Yulikayani *et al.*
Environmental Sciences Europe (2025) 37:203
https://doi.org/10.1186/s12302-025-01257-9

Environmental Sciences Europe

RESEARCH

Open Access

CONTRAST prioritisation tool: filtering and ranking contaminants of emerging concern in the marine environment using hazard-based approaches

Putu Yolanda Yulikayani[1,12], Bavo De Witte[1*], Aasim Ali[2], Jon Barber[3], Juan Bellas[4], Nicolas Briant[5], Steven Brooks[6], Are Sæle Bruvold[2], Megan Anne French[7], Ketil Hylland[2], Helen Kaberi[8], Víctor M. León[9], Samantha Martins[6], Aourell Mauffret[5], Francesca Molinari[10], Malcolm Reid[6], Joachim Sturve[10], Christof Van Poucke[11], Stig Valdersnes[2], Christophe Walgraeve[12], Kristof Demeestere[12] and Aina Charlotte Wennberg[6]

Abstract

Background Increasing numbers of chemicals with little-known adverse effects are released into the marine environment. The present study addresses the lack of marine-specific prioritisation schemes by developing a prioritisation tool for organic contaminants. This tool supports decision-making processes regarding which chemicals to study further in terms of their occurrences and biological effects in the marine environment. It was supported by a database containing approximately 1.13 million chemicals, developed within the PikMe project. Criteria for chemical prioritisation were identified by a comprehensive literature review, then selected using the outcomes of a survey among experts. The prioritisation tool consists of filtering chemicals in the PikMe database using three parallel schemes—persistence and bioaccumulation, toxicity, and persistence and mobility characteristics (step 1)—followed by scoring based on modes of action, occurrence, and emission (step 2) and ranking by the final score (step 3).

Results Around 8000 chemicals were selected by filtering (step 1). The top 100 resulted from step 3 comprises 6PPD as the highest-ranked compound and other chemicals with high diversity of uses, e.g. pharmaceuticals as the predominant category of use, industrial chemicals, personal care products, flame retardants, and plastic additives. These chemicals were ranked in the top 100 due to dominant influence of diverse prioritisation criteria.

Conclusions Using the hazard-based approach that encompasses different adverse effects that contaminants of emerging concern can exert, the marine-specific prioritisation tool can guide decision-making in monitoring, ecotoxicological studies, and regulations regarding contaminants of emerging concern in the marine environment.

Keywords Contaminants of emerging concern, Hazard-based assessment, Persistence, Bioaccumulation, Mobility, Toxicity

*Correspondence:
Bavo De Witte
Bavo.Dewitte@ilvo.vlaanderen.be
Full list of author information is available at the end of the article

Background

The advancement of technology has facilitated the synthesis and utilisation of an increasing number of chemicals, significantly enhancing convenience in daily life and promoting social development. Starting from around 1.6 million chemicals in 1970, there are now more than 279 million unique chemical substances, including organic and inorganic compounds, alloys, coordination compounds, minerals, mixtures, polymers, and salts registered in the Chemicals Abstracts Service (CAS) [1, 2]. In Europe, around 22,000 chemicals have been registered in the European Economic Area (EEA) market [3]. The extensive use of chemicals for diverse applications such as pharmaceuticals, plastic additives, pesticides (including biocides), personal care products (PCPs), and various industrial uses has resulted in increased prevalence of these chemicals in the marine environment [4].

Incomplete removal in wastewater treatment plants (WWTP), sewer leakage, direct discharge of untreated water from households and industries, runoff from farms, surface runoff, stormwater runoff, and atmospheric deposition are routes through which chemicals from land-based activities can enter the marine environment [5]. In addition, sea-based activities such as shipping, accidental spillage, mariculture, or emissions from antifouling paints contribute to the presence of chemical contaminants in the marine environment [6, 7]. While some chemicals are regulated and routinely monitored, most of the anthropogenic chemicals are unregulated, and their occurrences and (eco)toxicological effects are often not well-studied. These chemicals are referred to as contaminants of emerging concern (CECs) [8].

There is a need to improve knowledge on CECs in the marine environment in order to support regulatory agencies in reassessing their practices and considering CECs that are less routinely monitored. While the development of high-performance analytical instruments and new analytical approaches has made it possible to detect more compounds using semi-quantitative methods, such as non-target analysis or suspect screening, studies on the biological implications of lesser-known CECs is time consuming and has struggled to keep up with the number of new chemicals being detected in the environment [9, 10]. In addition, the development of quantitative tools to measure non-monitored CECs is needed for water monitoring strategies but is hindered by the wide diversity of physical–chemical properties of CECs [11]. Thus, prioritisation tools can help to focus on studying those chemicals that are likely to cause the most harmful effects, while improving water protection plans by adding prioritised substances to the list of monitored compounds.

In general, there are four existing approaches to prioritise CECs, i.e. prioritisation based on exposure, hazard, risk, or a combination of these approaches. The prioritisation approach based on exposure compares and ranks chemicals based on their potential environmental emissions, which can be calculated using production volume and characteristics that describe the condition of use from an environmental perspective [12, 13]. Hazard-based prioritisation assesses chemicals based on their persistence, bioaccumulation, and mobility characteristics, as well as acute and chronic toxicity including modes of action (MoA) (e.g. carcinogenicity, mutagenicity, endocrine disrupting potential) [14–16]. The risk-based approach has been the most widely reported in literature on prioritisation of CECs. The risk of a compound is defined as the ratio of exposure (predicted environmental concentration (PEC) or measured environmental concentration (MEC)) to effect (i.e. predicted no effect concentration (PNEC)) [17–19]. Some prioritisations have used a combination of hazard and risk [20], or exposure, hazard, and risk approaches [21], to more comprehensively assess potential impacts.

The behaviour and fate of CECs in the marine environment can be different compared with the freshwater environment due to differences in environmental conditions. The most notable dissimilarity is the higher salt content of the marine environment (NaCl concentration in saltwater is ~ 35 g/L, while in freshwater the concentration is < 1 g/L), which can lead to the salting out effect [22–24]. The salting out effects increases hydrophobicity and air–water partition coefficient of CECs, enhancing their partitioning into air, organic carbon, and lipid [25]. Moreover, the average pH of seawater is around 8.1–8.3 [26, 27], while freshwater has a wider variation of pH between 6.5 and 9.0 [28]. A difference in pH means different degrees of ionisation for weak acids and bases, which affects the sorption and bioavailability of compounds, mainly relevant for pharmaceuticals (e.g. citalopram, propranolol, carbamazepine) [26, 29]. Additionally, differences in sorption capacity between marine organic carbon (OC) and terrestrial OC have been observed and can alter the fluxes of hydrophobic organic compounds [30].

Considering the differences outlined above, a prioritisation of CECs that takes into account the specificity of the marine environment is likely to be different when compared with prioritisations that have been developed for the freshwater environment. The literature review performed in this study showed that 8 of 31 prioritisation schemes considered marine matrices, however, only four were developed specifically for selecting CECs in the marine environment. In addition, many lesser-known contaminants were not considered in these existing prioritisation schemes due to: (i) the use of a limited list of initial chemicals, (ii) the toxicity assessment that was only

based on acute and chronic toxicity and do not consider the compound potential MoA, (iii) the lack of occurrence data for most of the CECs. As a result, considerable amounts of CECs were neglected despite their potentially significant adverse effects.

The European project "Contaminants of Emerging Concern: An Integrated Approach for Assessing Impacts on the Marine Environment (CONTRAST)" has an overall objective to develop an integrated assessment framework involving effect-based monitoring tools to provide a holistic evaluation of the presence, impacts, and risks of CECs in the marine environment. Therefore, to select the most relevant CECs for the integrated assessment in the CONTRAST project, the present study has developed a prioritisation tool for organic CECs with a specific focus on the European marine environment [31]. This tool emphasis MoA in addition to acute and chronic toxicity values to comprehensively assess the toxicity and biological effects of chemicals. A literature review of existing prioritisation schemes was carried out to identify criteria used to prioritise compounds. The outcome of a survey on the criteria needed in a marine-specific prioritisation were used as a basis to develop a prioritisation tool that is useful and adequate for selecting CECs in the marine environment and effective when tested against compounds from existing priority lists.

Materials and methods

Review on existing prioritisation schemes

A list of existing prioritisation schemes was compiled by searching the Web of Science™ database (Feb–Mar 2024) for studies published between 2011 and 2024 using the combination of keywords related to "prioritisation" or "prioritization", "chemicals", "contaminants of emerging concern" or "organic contaminants", and "environment". Studies were selected based on titles and abstracts, focusing on schemes addressing environmental (marine, freshwater, terrestrial) impacts of chemicals, excluding those focused solely on human health. Only studies that proposed a prioritised chemical list and described their selection method in the "Materials and methods" section were included. Additional schemes were also sourced from environmental agencies, monitoring networks, and regional sea conventions.

Subsequently, a more detailed examination was carried out to extract essential information from each prioritisation schemes. This information was compiled into a table and was used as the starting point for the selection of prioritisation criteria in this study. The information contained:

- Geographical context: the region where the prioritisation scheme was applied.
- Assessed compounds: the initial list of chemicals assessed using the prioritisation scheme (list of chemicals from literature and reports or list of chemicals detected from a monitoring program).
- Matrix: the type of matrix considered (e.g. freshwater, sea water, drinking water, air, and/or sediment).
- Criteria: the chemical characteristics used to prioritise compounds (e.g. persistence, toxicity, mobility, risk, and emission).
- Additional considerations: evaluation outside the aspect of chemical characteristics (e.g. criteria weighting, data source prioritisation, and data gaps).
- Prioritisation method: how compounds were prioritised (categorisation or ranking).

Selection of prioritisation criteria

An online survey was conducted among project participants from various institutes (SI 1.1) to gather expert input on criteria for the CONTRAST prioritisation tool. Completed by 18 respondents between 4 and 29 March 2024 (SI 1.2), the Google Forms survey asked participants to rate their agreement with including specific criteria—identified from a review of existing schemes—on a 5-point Likert scale (1 = Totally disagree to 5 = Totally agree). Respondents were also asked to suggest additional modes of action (MoA) and rank each criterion by importance. A criterion was included in the tool if more than 50% of respondents agreed or totally agreed. The analysis of the survey data was performed by plotting Likert scales using package 'likert' [32] in R (version 4.3.2) [33].

Development of a prioritisation tool

General overview and use of the PikMe tool

An overview of the prioritisation process is illustrated in Fig. 1. The prioritisation of CECs was assisted using the PikMe tool (preliminary version of 5 September 2024), described in more details in Wennberg et al. 2025 [34]. The PikMe tool was developed by The Norwegian Institute for Water Research (NIVA) and The Climate and Environmental Research Institute (NILU), and comprises a database of 1.13 million compounds chemicals gathered from the European Chemicals Agency (ECHA), CompTox Chemical Dashboard of the United States Environmental Protection Agency, the NORMAN Network substance database, quantitative structure–activity relationship (QSAR) predictions using Open (Quantitative) Structure–activity/property Relationship App (OPERA), and Estimation Program Interface (EPI) Suite predictions.

The selection of compounds began by excluding those regulated in the European Union, followed by filtering

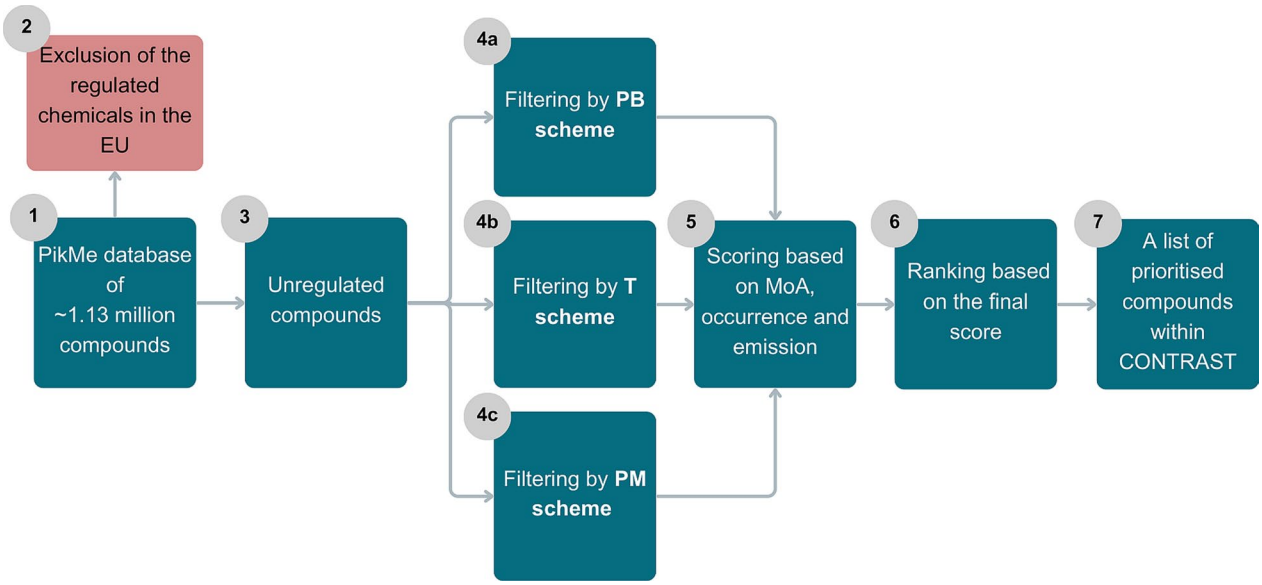

Fig. 1 Overview of the CONTRAST prioritisation tool for filtering using persistence and bioaccumulation (PB), toxicity (T), and persistence and mobility (PM) schemes and scoring based on modes of action (MoA), occurrence, and emission

through three parallel schemes—persistence and bioaccumulation (PB), toxicity (T), and persistence and mobility (PM)—using the PikMe tool. PikMe transparently assessed each chemical's persistence (P), bioaccumulation (B), mobility (M), and toxicity (T) using diverse data sources. Filtered compounds were scored based on mode of action, presence in monitoring databases, and emissions, then merged and ranked to produce a prioritised list. The CONTRAST prioritisation focused on organic compounds via PB and PM schemes, while the T scheme could include other types (e.g. metals, organometallics, inorganics), though these were not the main focus.

Filtering of compounds (step 1)

First, the chemicals currently regulated or not approved in the EU were excluded to focus on CECs. The exclusion of regulated compounds was made possible using the PikMe tool, which includes in its database lists of hazardous compounds from European Regulations (downloaded in January 2024, SI 1.3). In addition to regulated compounds included in PikMe, a list of active substances in plant protection products currently not approved for use in the EU was gathered from the "EU Pesticide Database > Active Substances" [35] to exclude more substances that are currently not in use. Altogether, there were 7193 regulated compounds excluded from the CONTRAST prioritisation scheme. The exclusion of regulated compounds was carried out using CAS numbers. Thus, compounds without clear CAS numbers (e.g. mixtures) cannot be excluded from the present prioritisation despite their presence in the lists of EU regulated compounds.

The filtering was carried out using the three parallel schemes (PB, T, and PM), which generated three respective lists of compounds. For each approach, cut-off values were used to filter the compounds (Table 1). The cut-off values for persistence, bioaccumulation and toxicity were adopted with few modifications from the ECHA's Guidance document on assessing persistent, bioaccumulative, and toxic (PBT)/very persistent and very bioaccumulative (vPvB) [36], while the cut-off value used for mobility was adopted from the Classification, Labelling, and Packaging (CLP) Regulation (EC 1272/2008) [37]. A cut-off value of log $K_{OW} > 3$ was used to assess bioaccumulation potential, diverging from ECHA's threshold of log $K_{OW} > 4.5$. This adjustment was made to acknowledge the potential for reduced solubility resulting from salting out effect in the marine environment. As compounds' solubility decreases in seawater, their affinity to lipids increases, leading to a higher potential for bioaccumulation in marine organisms [24]. In addition, Regulation (EC) No 1907/2006 mentioned that as part of the information requirements of Annex IX of REACH, the study of bioaccumulation in aquatic species, preferably fish, should be conducted if log $K_{OW} > 3$ [38].

When filtering chemicals based on toxicity, a selection based on MoA was also applied, alongside the selection based on acute and chronic toxicity adopted from the ECHA PBT/vPvB assessment [36]. The MoA used for filtering comprised developmental toxicity,

Table 1 Cut-off values used to filter the compounds based on persistence and bioaccumulation, toxicity, and persistence and mobility

Scheme	Cut-off value	Reference	Score
Persistence and bioaccumulation	Half-life in water and sediment in freshwater, estuarine, and marine environment > 40 days OR QSAR not readily biodegradable AND	[36]	PB overall score = (CONTRAST P score x certainty score) + (CONTRAST B score x certainty score)
	Log K_{OW} > 3		
Toxicity	Acute toxicity L(E)C_{50} < 100 µg/L OR Chronic toxicity EC_{10} or No Observed Effect Concentration (NOEC) < 100 µg/L	[36]	$T_{ecotoxicity}$ score = CONTRAST T score x certainty score T_{MoA} score 1 if there is only one mode of action, 2 if there are multiple MoA T overall score = Highest score between $T_{ecotoxicity}$ and T_{MoA}
	Developmental toxicity (TEST) > 0.5 OR Ames mutagenicity (TEST) > 0.5 OR endocrine disruption estimated as active (agonist, antagonist, or binding) by CERAPP or CoMPARA	[16]	
Persistence and mobility	Half-life > 40 days OR QSAR not readily biodegradable AND	[36, 37]	PM overall score = (CONTRAST P score x certainty score) + (CONTRAST M score x certainty score)
	Log Koc < 3 (for neutral compounds) OR Log Koc for pH 4–9 < 3 (for ionisable compounds)		

mutagenicity, and endocrine-disrupting properties whose prediction data are included in PikMe database. Developmental toxicity and mutagenicity data were predicted using QSAR models included in the Toxicity Estimation Software Tool (TEST) [39]. The QSAR model for developmental toxicity was developed by the Computer Assisted Evaluation of industrial chemical Substances According to Regulations (CAESAR) project, using existing human and animal data on potentially teratogenic substances [40]. The Ames test, which uses histidine-dependent *Salmonella typhimurium* strains to identify mutagenic compounds, has been widely used as amethod for assessing mutagenicity. To avoid in vitro testing, a QSAR model was developed using a benchmark dataset of ~ 6,500 compounds to predict Ames test outcomes [41]. Endocrine disrupting potential was assessed using oestrogen receptor activity and androgen receptor activity, predicted using models developed by the Collaborative Estrogen Receptor Activity Prediction Project (CERAPP) and the Collaborative Modelling Project for Androgen Receptor Activity (CoMPARA) [42, 43].

A scoring system for persistence, bioaccumulation, mobility, and environmental toxicity was also implemented in PikMe for filtering or ranking chemicals according to the risk of having each of these properties (SI 1.4). Since the criteria for potentially persistent, mobile, and toxic on the PikMe tool were similar to this study's cut-off values, the PikMe score for these properties could be used directly to filter unregulated compounds in the PikMe database. As for bioaccumulation, the cut-off value of Log K_{OW} > 3 was set on the PikMe tool to select potentially bioaccumulative compounds.

There were still hundreds of thousands of chemicals being filtered using the cut-off values of each scheme. Therefore, to further reduce to a manageable number of compounds, two different approaches were applied in parallel. The first approach was to retain only compounds that have occurrence data. Filtering based on the existence of occurrence data was possible as the PikMe tool also includes a list of compounds present in the NORMAN EMPODAT database [44]. This database contains occurrence data of CECs that are not necessarily included in regular monitoring programs but present in the environment (i.e. water, sediment, biota, suspended particulate matter, soil, sewage sludge, and air). The occurrence data were obtained from research projects, national monitoring program, etc. and were gathered across Europe and beyond [21]. Within the CONTRAST prioritisation, both marine and freshwater data were taken into account.

The second approach to reduce the number of compounds was to retain those with reliable data using the reliability score in PikMe. The reliability score indicates the type of data (experimental or estimated data), and the reliability of the estimated data used to determine persistence, bioaccumulation, mobility, and environmental toxicity scores. Experimental data were given a reliability score of 2, while estimated data were given a score between 0 and 1, which represented the applicability domain (AD) index. The AD of a model is defined as a multidimensional chemical space of its training set, encompassing not only chemical structures and physicochemical properties but also mechanistic insights and the metabolic domain relevant to the modelled phenomenon [45]. If a chemical falls within the AD of the training set

chemical space, then the prediction can be considered reliable [46]. Using this approach, the compounds with a reliability score ≥ 0.8 were retained, indicating that they have experimental data or reliable predictions regarding their persistence, bioaccumulation, toxicity, and mobility.

Following the reduction in the number of filtered compounds using the two parallel approaches, 1758 compounds were filtered by the PB scheme, 7585 compounds by the T scheme, and 1071 compounds by the PM scheme. The parameters set in PikMe for filtering are detailed in SI 1.5. The compounds filtered by each scheme were assigned a score based on their persistence, bioaccumulation, toxicity, and mobility characteristic using a decision tree, elaborated in Fig. 2–4.

In the PB scheme (Fig. 2), compounds were scored based on persistence (P/vP) and bioaccumulation (B/vB) criteria. If these were not met, compounds were labeled as potentially persistent and/or bioaccumulative. As the ECHA Guidance for PBT assessment considers compounds with a Log $K_{OW} > 4.5$ as potentially bioaccumulative, which is less conservative than the cut-off value of Log $K_{OW} > 3$ for potentially bioaccumulative compounds

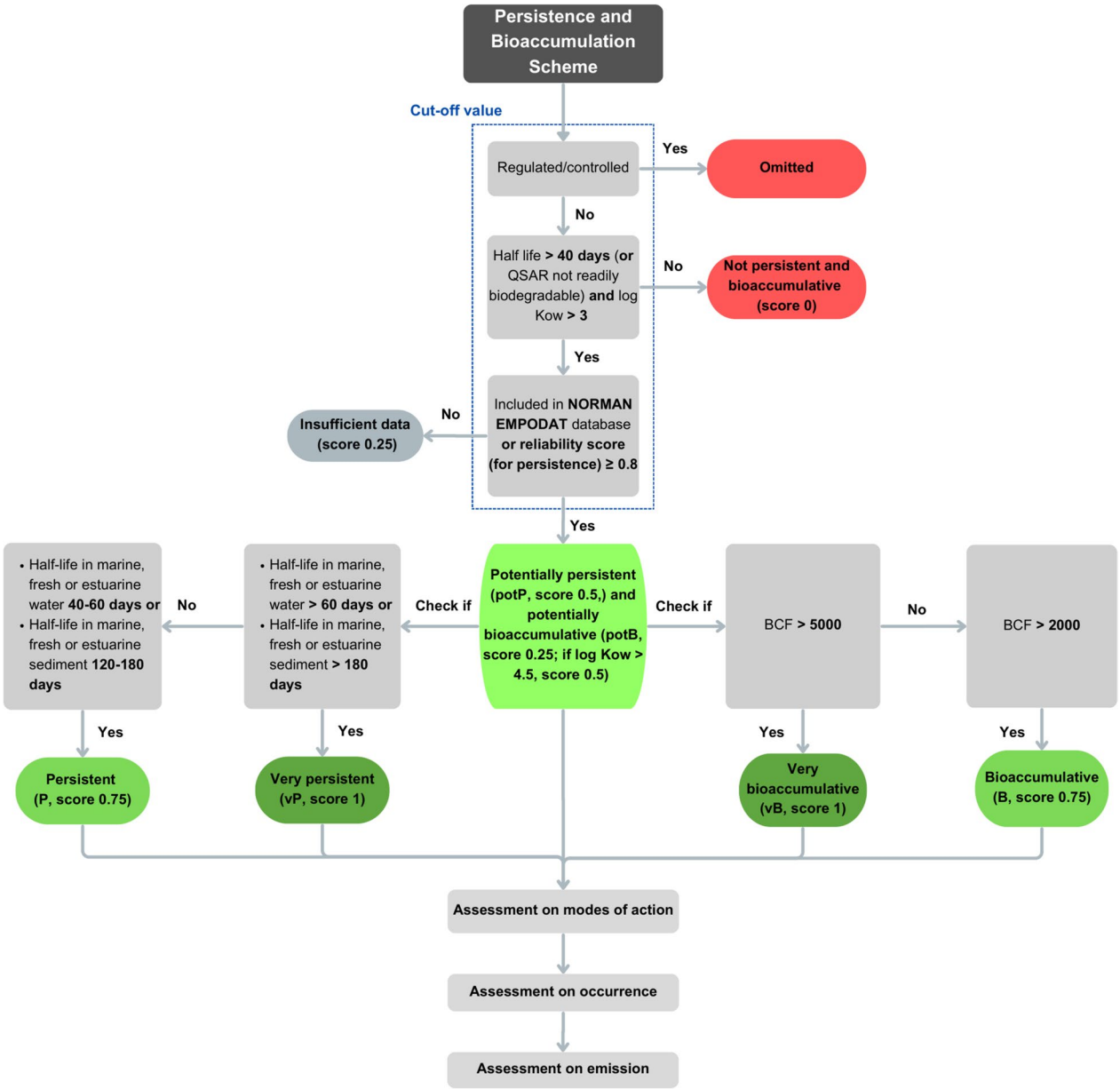

Fig. 2 CONTRAST prioritisation scheme based on persistence and bioaccumulation. (BCF: bioconcentration factor)

used in our prioritisation scheme, a higher score was given to compounds with a Log $K_{OW} > 4.5$. In the toxicity-based scheme (Fig. 3), compounds scored higher if they showed acute toxicity ($L(E)C_{50} < 10$ µg/L), chronic toxicity (EC_{10} or NOEC < 10 µg/L), or multiple modes of action. Others were marked as potentially impactful (potI). Due to limited marine toxicity data, non-marine species data were also considered. In the PM scheme (Fig. 4), compounds with Log $K_{OC} < 2$ at any pH (or pH 4–9 for ionisable compounds) were deemed very mobile; others were considered mobile with lower scores. Persistence was further assessed via half-life data.

Due to varying data types and quality across the three schemes (Figs. 2, 3, 4), a system was developed to distinguish compounds assessed with experimental data from those with predicted data. Certainty scores, based on the PikMe reliability score, were used to weight the PB, T, and PM scores. Experimental data (PikMe score

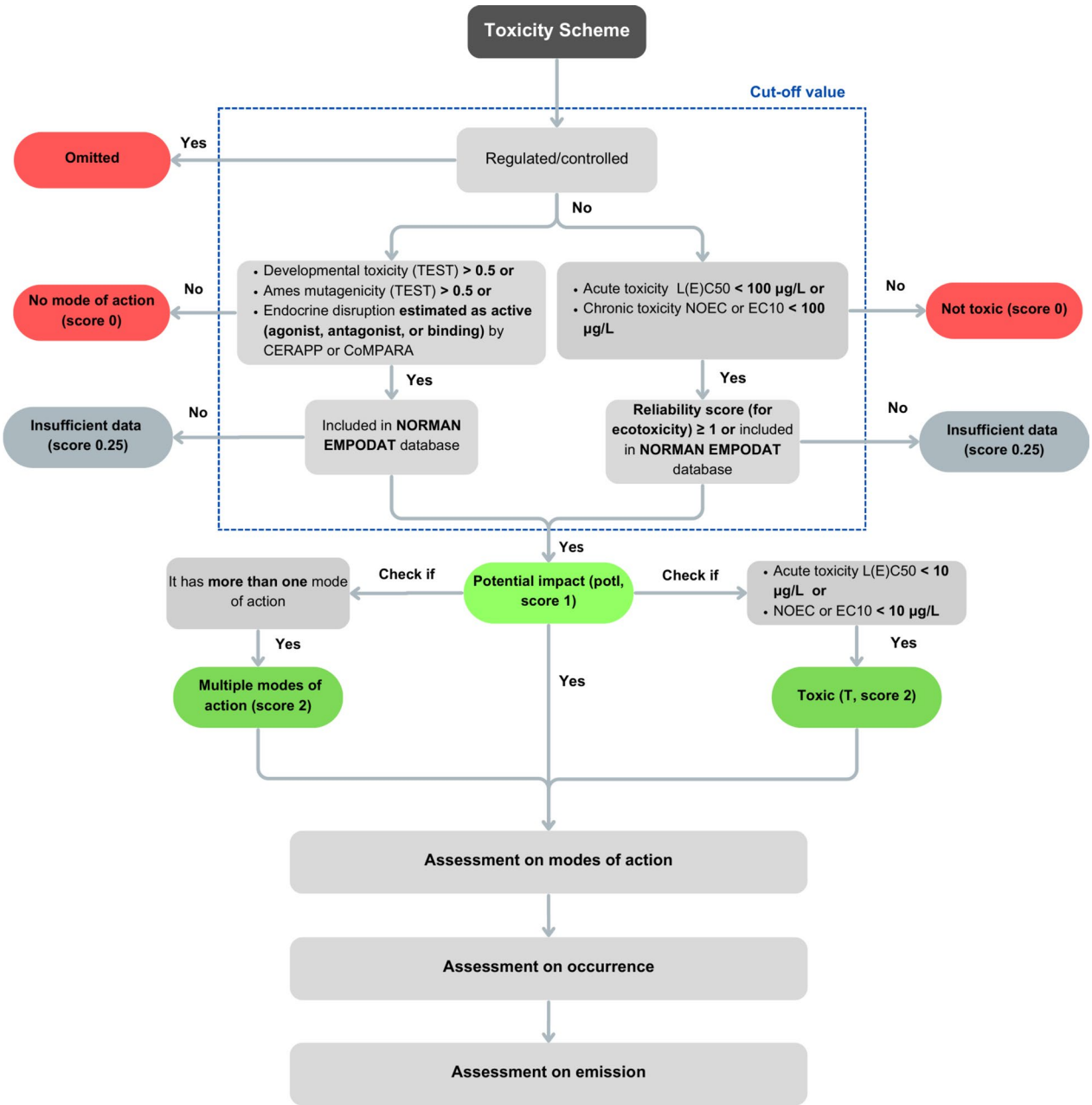

Fig. 3 CONTRAST prioritisation scheme based on toxicity

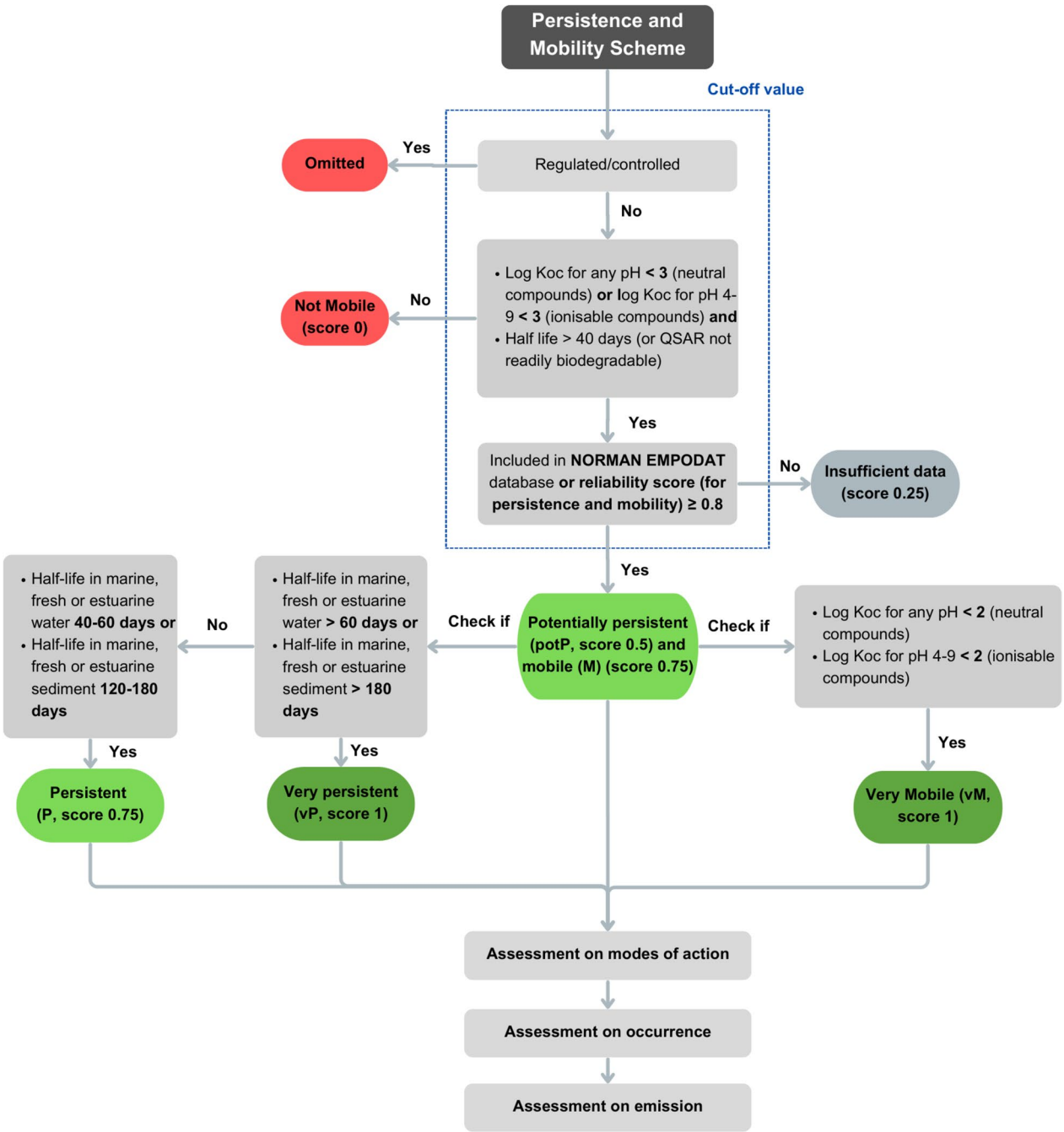

Fig. 4 CONTRAST prioritisation scheme based on persistence and mobility

2) received higher weights, while predicted data (PikMe or AD index 0–1) received lower weights. The certainty score was calculated as PikMe score divided by two, with a minimum score of 0.05 for compounds lacking reliability data. Final scores were obtained by multiplying the CONTRAST prioritisation scores by the corresponding certainty score (0.05–1) (Table 1).

Assessment of MoA, occurrence, and emission (step 2)

After filtering and scoring using the prioritisation schemes, the selected compounds were assessed and scored based on their MoA, occurrence, and emission. In the assessment of MoA, six parameters which consist of carcinogenicity, mutagenicity, reproductive toxicity, specific target organ toxicity, endocrine disrupting potential,

and developmental toxicity were evaluated using data gathered from data sources detailed in SI 1.6 Table 1. For each data source, dedicated criteria were applied to determine whether a compound could be considered to have a certain MoA. Based on the type of data available for MoA assessment, a compound received a score for each parameter (SI 1.6 Table 2).

Some compounds are assigned a certain MoA in multiple data sources. In this case, the highest priority and score was given to the regulatory framework data, i.e. REACH data, which categorise compounds as confirmed or suspected of having the MoA in question. A confirmed MoA was assigned a higher score than a suspected MoA. Based on reliability, data measured in vitro or in vivo received a higher score than estimated data. If there was no data available, the compound received a score of 0.1. Otherwise, it was assigned a score of 0 if the compound was examined and classified as not having the MoA in question.

The occurrence of each compound was assessed based on the presence or absence in the NORMAN EMPODAT database [44], exclusively in marine matrices, and in the Norwegian Monitoring data from the Norwegian Environmental Agency database Vannmiljø [47], to which this study has access. The presence/absence approach gives equal weight to CECs with limited occurrence data compared with widely studied CECs with much occurrence data. The matrices considered in the Norwegian Monitoring data were air, sediment, water, and biota both from the marine and freshwater environment in Norway. As for the scoring, the presence > LOQ (Limit of Quantification) was given a higher score than the presence > LOD (Limit of Detection) (SI 1.7), as presence > LOQ indicates not only that a compound can be detected but also quantified within certain limit of confidence.

To assess the emission of compounds, data regarding the production and use patterns were used, which was accessed from the ECHA and ChemExpo Knowledgebase websites. The method used to assess the emission was adopted from the NORMAN Prioritisation Methodology [48]. Compounds with higher production volume are more likely to be present in the environment and were, therefore, given higher scores (SI 1.8 Table 1). The use pattern reflects the activities that contribute to the use of a substance from the worker's perspective and the environmental perspective. According to the NORMAN Prioritisation Methodology, there are four types of use patterns:

- Used in the environment: direct releases to the environment (e.g. biocides and pesticides used to protect crops or UV filters in sunscreens from bathers);
- Wide dispersive use: dispersed source releases to the environment (e.g. substances released from WWTPs such as pharmaceuticals, PCPs, flame retardants, and plastic additives);
- Non-dispersive use: identified and controlled releases from small numbers of point sources;
- Controlled system: no direct release to the environment (i.e. substances used in the controlled process in industry or used as intermediate in a closed system).

The use pattern was evaluated based on the data of 130 function categories, each consisting of a list of compounds, found on the website "ChemExpo Knowledgebase > Chemical Function Categories" [49, 50]. Each function category (e.g. biocide, UV stabiliser, flame retardant) was then classified to an appropriate use pattern and assigned to a certain score (SI 1.8 Table 2 and 3).

Since this study focuses on CECs in the marine environment, an emission assessment was conducted to identify compounds with potential sea-based sources. This was supported by a list from Tornero and Hanke (2016), which compiled marine contaminants from sea-based sources in Europe, excluding those from atmospheric transport [6]. The authors focused on the European context and obtained the information from regulatory and RSCs, literature, reports, assessments, and research projects. Using this list, the chemicals with potential sea-based sources were given an additional score (SI 1.8 Table 4). The production score and

Table 2 Sensitivity analysis of the CONTRAST prioritisation tool to final scores and rank of compounds in the top 100 priority list

Scenario id	Modified assessment	Mean absolute deviation	Mean relative deviation (%)	Proportion of compound with rank change ≤ 20 positions (%)
Scenario 1	PB	– 0.26	– 5.28	72
Scenario 2	PM	– 0.08	– 1.66	91
Scenario 3	T	– 0.26	– 5.28	72
Scenario 4	MoA	– 0.27	– 5.50	86
Scenario 5	Occurrence	– -0.66	– 12.66	56
Scenario 6	Emission	– 0.25	– 5.18	83

Table 3 Number and percentage of chemicals in the list of compounds for validation (n = 290) selected by using the PB, PM, and/or T schemes in the CONTRAST prioritisation tool

Scheme(s)	Approach 1[a] (number; percentage (%))			Approach 2[b] (number; percentage (%))		
	Positive controls (n = 267)	Negative controls (n = 34)	Unknowns (n = 34)	Positive controls (n = 267)	Negative controls (n = 34)	Unknowns (n = 34)
PB	108; 40.4	0; 0.0	2; 5.9	24; 9.0	0; 0.0	0; 0.0
PM	50; 18.7	1; 2.9	5; 14.7	17; 6.4	1; 2.9	1; 2.9
T	217; 81.3	3; 8.8	10; 29.4	62; 23.2	0; 0.0	3; 8.8
PB and PM	22; 8.2	0; 0.0	2; 5.9	5; 1.9	0; 0.0	0; 0.0
PB and T	104; 38.9	0; 0.0	2; 5.9	23; 8.6	0; 0.0	0; 0.0
PM and T	44; 16.5	0; 0.0	5; 14.7	17; 6.4	0; 0.0	1; 2.9
PB, PM, and T	19; 7.1	0; 0.0	2; 5.9	5; 1.9	0; 0.0	0; 0.0
PB, PM, or T	224; 83.9	4; 11.8	10; 29	63; 23.6	1; 2.9	3; 8.8

[a] Applying the CONTRAST prioritisation tool while including EU regulated compounds

[b] Applying the CONTRAST prioritisation tool while excluding EU regulated compounds

use pattern score were grouped, while the sea-based sources score was kept separately to give it more weight in the overall emission score (SI 1.8 Eq. 1).

Final scoring of compounds (step 3)

Following the assessment of MoA, occurrence, and emission, a final score was calculated based on the individual score for each assessment and the overall score from the PB scheme, the PM scheme, and the T scheme (Eq. 1). Subsequently, all the compounds selected and scored using the prioritisation tool were combined and ranked based on their final score.

$$\text{Final score} = \text{PB} + \text{T} + \text{PM} + \text{MoA} + \text{Occurrence} + \text{Emission} \quad (1)$$

Validation of the CONTRAST prioritisation tool

As a way to check the effectiveness of the CONTRAST prioritisation tool, an assessment was carried out on chemical contaminants prioritised under different European regulatory frameworks and RSCs, compiled in a report from Tornero and Hanke (2016) on the identification of marine chemical contaminants released from sea-based and non-sea-based sources [51]. This compilation was used as a list of compounds for validation, which also including the compounds in the WFD 1st [52], 2nd [53], 3rd [54], 4th [55], 5th [56] Watch Lists. The details on the lists of regulations used for the validation was provided in SI 1.3. As the CONTRAST prioritisation tool was developed to select and prioritise organic compounds, metals/metalloids, organometallic compounds, and inorganic compounds were excluded from this validation process.

The compounds prioritised under WFD, EU Biocide Regulation, Commission Regulation (EU) No 37/2010, EMSA Dispersant Inventory and RSCs were considered relevant as they have been prioritised due to their adverse effects on the environment. Hence, they functioned as "positive controls", consisting of 267 compounds. Other regulated compounds in the report of Tornero and Hanke (2016) (e.g. chemical weapons and feed additives) were not necessarily linked to environmental adverse effects, and thus, served as a subset of "unknown" (34 compounds). The list of compounds for validation was completed with a subset of "negative controls", comprising compounds that should not be selected and prioritised due to their low negative impacts on the environment. These compounds were present in the OSPAR Agreement 2013–11 regarding the list of substances used and discharged offshore which are considered to pose little or no risk to the environment (PLONOR) [57], from which 34 organic compounds were gathered. In total, there were 335 organic compounds in the list of compounds for validation.

This validation was performed using two different approaches: (1) applying the CONTRAST prioritisation tool, without excluding EU regulated compounds (see section *Filtering of compounds (step 1)* in Materials and Methods), and (2) applying the prioritisation tool, excluding the regulated compounds. By performing the validation using the first approach, the effectiveness of the CONTRAST prioritisation tool to select relevant compounds (i.e. compounds under regulation for environmental protection) was evaluated. Therefore, it was expected that most of the positive controls would be selected, whilst the opposite was expected when applying the second approach as the aim was to check the effectiveness of the exclusion of EU regulated compounds in the selection of CECs (i.e. unregulated compounds). The negative controls were expected not to be selected or ranked low in both approaches.

Dimensionality reduction and clustering of top 100 priority compounds

To explore feature similarities among the top 100 ranked compounds from the CONTRAST prioritisation, dimensionality reduction techniques were applied to simplify the dataset, which included variables such as PB, T, PM, MoA, occurrence, and emission score. Both linear (principal component analysis (PCA)) and non-linear methods ((t-distributed stochastic neighbour embedding (t-SNE) and uniform manifold approximation and projection (UMAP)) were used. PCA was first applied to assess linear relationships; if non-linear patterns dominated, t-SNE and UMAP were used, and the results were compared. Prior to applying t-SNE and UMAP, the data were scaled. Clustering was performed after dimensionality reduction using two unsupervised methods for comparison: k-means and Gaussian mixture models (GMM). Cluster analysis is an unsupervised technique that reveals natural grouping based on the similarity of samples [58].

Prior to clustering with k-means, the number of clusters (k) was determined using the elbow method (SI 1.9 Fig. 2). A point where the rate of decrease in the within-cluster sum of squares (WCSS) begins to stabilise indicates the optimal number of clusters [59]. Regarding GMM, the optimal number of components and clusters was determined using Bayesian information criterion (BIC), which is one of the most common approaches to determine the number of components and clusters in mixture modelling [60]. The clustering performance was then evaluated using the silhouette index (S(i)) [61], ranging from -1 to 1, where 1 indicates perfect clustering, 0 indicates intersecting clusters, and -1 indicates misclassifications. The dimensionality reduction and clustering analysis was carried out using R (version 4.3.2) [33].

Sensitivity analysis of the top 100 priority compounds

A sensitivity analysis was performed on the compounds filtered using PB, PM, and T schemes to evaluate the impact of each of the six individual assessments in the CONTRAST prioritisation (i.e. PB, PM, T, MoA, occurrence, and emission) on the final score and rank of the retained compounds. Using a one-at-a-time approach, each assessment's scoring was varied individually while others remained at default, resulting in six scenarios (SI 1.10). The deviations from the baseline score of each compounds, expressed as absolute deviation and relative deviation (Eqs. 2 and 3 respectively), and the changes in rank were observed to identify the most influential assessment. To compared the robustness of each compound to different scenarios, the mean and standard deviation of the deviation from the baseline final score was calculated. Furthermore, the percentage of compounds in the top 100 with a rank change (Eq. 4) of ≤ 20 positions was also evaluated for each scenario.

$$\text{Absolute deviation} = \text{Final score of the scenario} - \text{baseline score} \quad (2)$$

$$\text{Relative deviation} = \left(\text{Final score of the scenario} - \text{baseline score}\right)/\text{baseline score} \quad (3)$$

$$\text{Rank change} = \text{Scenario rank} - \text{baseline rank} \quad (4)$$

Results

Evaluation of existing prioritisation schemes

In total, 31 prioritisation schemes were selected and examined in detail to extract the key information (summarised in SI 1.11 Table 1). Most of the prioritisation schemes were developed and implemented in Europe, while 11 of the inventoried prioritisation schemes were implemented in Asia or North America. More than half of the prioritisation schemes utilised existing lists of compounds from literature or reports, such as the United States Environmental Protection Agency (EPA) Priority Pollutants List under the Clean Water Act [14], OSPAR's Lists of Chemicals for Priority Action and List of Substances of Possible Concern [20], or the Artic Monitoring and Assessment Program (AMAP) database of chemicals [62]. Other schemes utilised lists of chemicals detected within monitoring programs at specific locations [11, 63], whilst some used both lists of chemicals from literature and monitoring programs [9, 64].

As for the matrices considered in existing prioritisation schemes, surface water and sediment were the most considered, followed by biota (SI 1.11 Fig. 1). From eight prioritisation schemes that considered marine matrices, only four were specifically developed for the marine environment. One of those prioritisation schemes was developed within the OSPAR Convention, which aims to protect the marine environment of the North-East Atlantic [65]. Some prioritisation schemes considered other matrices such as drinking water [66], groundwater [67], WWTP effluent [8], and suspended particulate matter [68].

There were 17 criteria identified from the selected prioritisation schemes that were used to assess and prioritise CECs (SI 1.11 Fig. 2). Persistence, bioaccumulation, and toxicity were the most used criteria in the existing prioritisation schemes. To assess the persistence of compounds, the degradation half-life was generally used as a criterion [20, 69]. As for the compounds without available half-life data, biodegradation screening test results from the Organization for Economic and Co-operation and Development technical guidance (OECD TG) or

QSAR prediction as "not readily biodegradable" were considered as an indication of potentially persistent compounds [70]. Log K_{OW} and bioconcentration factor (BCF) were used to determine bioaccumulative and potentially bioaccumulative compounds [71, 72]. The Log K_{OA} (octanol–air partition coefficient) was applied when bioaccumulation in air-breathing animals was considered [16, 73]. The occurrence and detection frequency were also commonly considered, which can be particularly relevant for CECs because they can be present in the marine environment not only due to their persistence but also to their continuous discharge (pseudo-persistence) [74].

The toxicities of compounds were evaluated using acute toxicity (e.g. LC_{50} or EC_{50}) and/or chronic toxicity data (e.g. EC_{10}, NOEC, or chronic value (ChV), which is the geometric mean of NOEC and lowest observed effect level (LOEC)). Generally, there was no distinction between the use of ecotoxicity data for freshwater or marine organisms in the existing prioritisation schemes. The criteria on the most reported MoA of chemicals (i.e. endocrine disrupting potential, carcinogenicity, mutagenicity, reproductive toxicity) were generally assessed using data from Article 59 (candidate list), Annex XIV (authorisation list), Annex XVII (restriction list) of the REACH regulation (EC 1907/2006); the Community Rolling Action (CoRAP) list of the ECHA; and the CLP regulation (EC 1272/2008) [15, 18]. Other less reported MoA were also considered in some prioritisation schemes, such as xenobiotic metabolism-related effects [69], developmental toxicity, and neurotoxicity [16].

Many of the identified prioritisation schemes utilised risk-based approaches to prioritise CECs. Some of them used more specific risk-based criteria, such as the frequency of exceedance and the extent of exceedance. The frequency of exceedance addresses the spatial distribution of the exposure of biota at concentrations above a certain effect threshold, i.e. PNEC [68, 69], toxicity quotients (TQs), and exposure-activity ratios [75]. While this criterion indicates that the compounds are widely distributed, it cannot differentiate between compounds present at low concentrations close to their effect threshold and those that occurred at the highest concentrations. To address this, some prioritisation schemes have also used the extent of exceedance as a criterion [18, 21, 68].

Tonnage and emission are two criteria used in several prioritisation schemes to identify chemicals that are likely to occur in the environment. Tonnage only addresses the production volume, while emission also takes into account the usage characteristics of a chemical. This parameter describes specific information on chemicals' usage conditions linked to the possibility of emission [76]. The prioritisation schemes that included tonnage or emission mostly used REACH data on tonnage and environmental release category (ERC) to evaluate emission [12, 66]. Some used the Substances in Preparations in Nordic Countries (SPIN) database as source of information on annual import and production in the Nordic Countries and on the exposure estimation to five different recipients consisting of surface water, soil, air, sewage treatment plants, and consumers [15, 77].

The ionisability of chemicals was considered in several prioritisation schemes because the bioavailability and toxicity of ionisable chemicals depend on water pH. One method to evaluate ionisability is to check whether more than 10% of the molecules of a chemical are neutral at typical freshwater pH [71]. Neutral compounds are more likely to be absorbed by aquatic organisms due to the increased likelihood of crossing cell membranes [78]. A QSAR-based study that classified organic compounds with potential persistence and mobility characteristics (PMOCs) by charge and ionisability showed that ionic compounds tend to be more persistent and mobile than ionisable and neutral compounds [70].

In addition to the above criteria, some existing prioritisation schemes considered other aspects related to data availability, data source, and the importance of certain criteria compared with others (SI 1.10 Fig. 3). For many compounds, there is a lack of reliable toxicity and/or occurrence data for risk assessment and prioritisation. To address this lack of information, several of the identified prioritisation schemes grouped compounds into action categories based on the sufficiency of monitoring effort, analytical capabilities for quantification, and experimental toxicity data. By using this approach, substances with insufficient data would not be disregarded and knowledge gaps could be identified. Therefore, specific actions to fill the missing data gaps could be taken. These actions include increasing monitoring efforts for compounds with a lack of occurrence data, improving insufficient analytical capabilities for quantification, or performing effect assessment for compounds with scarce toxicity data [8, 21, 68].

In some prioritisation schemes, data source was prioritised to account for the varying quality of data generated from different sources. Generally, the data from REACH dossiers were given the highest priority [15, 70]. Additionally, data derived from experiments were prioritised over predictions from structure–activity relationship models [69]. Criteria weighting was used to emphasise the relevance of certain criteria based on monitoring data (i.e. frequency of exceedance, extent of exceedance, and spatial distribution) [18] or the environmental risk from ecotoxicological perspective [64].

Following assessments using the criteria mentioned above, compounds were prioritised using different approaches, which consist of categorisation, ranking, and

filtering. Using the categorisation approach, compounds are grouped based on the degree of priority [72, 75], action categories [21], or hazard (e.g. persistence, mobility, toxicity, bioaccumulation) characteristics [14, 79]. Categorisation was performed using a set of criteria for each category [8, 66] or using a score [69, 71]. The compounds in each category were considered equal in most of the schemes that used categorisation as a prioritisation approach. However, some prioritisation schemes ranked the compounds in each category using a score [18, 68]. Most prioritisation schemes used the score to rank or prioritise compounds while some also used it to filter and reduce the number of compounds in the final priority list [15, 80, 81].

Selection of priority compounds using the CONTRAST prioritisation tool

Based on the results of the survey to select appropriate prioritisation criteria for the CONTRAST prioritisation tool, the filtering and ranking approaches were developed to select CECs (SI 1.12). As a first step, 1.13 million compounds in the PikMe database were filtered using the PB, T, and PM schemes in parallel. Filtering using the PB scheme resulted in 1758 compounds being selected, while T and PM schemes selected 1071 and 7585 compounds, respectively. In total, 8548 compounds were ranked after assessing the compounds in the PB, T, and PM lists for further assessment on MoA, occurrence, and emission and merging them into one list (hereafter referred as the CONTRAST priority list, SI 2.1).

The top 100 priority compounds were extracted from the CONTRAST priority list, and will be used as the basis for selecting the compounds to be further studied in the CONTRAST project, focusing on environmental fate, distribution, and effects. The top 100 priority compounds consisted of compounds with a high diversity of use categories. To better illustrate the diversity of use, the top 100 priority compounds were categorised into 10 general use categories based on their assigned chemical classes in PubChem [82], including flame retardants, hormones, industrial chemicals, metals, polycyclic aromatic hydrocarbons (PAHs), PCPs, pesticides, per- and polyfluoroalkyl substances (PFAS), pharmaceuticals, and plastic additives. Using the treemap chart, each use category was represented with a rectangle, whose area was proportional to the number of compounds assigned to the given use category (Fig. 5). The top 100 priority compounds were dominated by pharmaceuticals followed by industrial chemicals.

UMAP with k-means clustering of the top 100 priority compounds, which was the most suitable dimensionality reduction and clustering (SI 1.9), revealed four clusters, consisting of cluster 2 and 3 which were well distanced from cluster 1 and 4. Cluster 1, the largest, comprising almost half of the top 100 priority compounds and characterised by maximum T score (T score = 2) and mostly PB and occurrence score of > 0.5, indicating PBT compounds that were present in the environment. Cluster 2 (15 compounds) also had maximum T score but none had both PM and PB score of 0, suggesting that they were either persistent and mobile or persistent and bioaccumulative. Cluster 3 grouped compounds with high occurrence scores (≥ 1) and moderate toxicity (T score = 1). Cluster 4 included highly toxic compounds (T score = 2) with MoA scores above 0.5, indicating strong biological effects. Full details, including scores, use categories, chemical classes, and cluster assignments, are provided in SI 2.2.

Sensitivity analysis of the top 100 priority compounds revealed predominantly negative absolute and relative deviations, mostly within -10%. This indicates that final scores were generally lower than the baseline and that the results were relatively insensitive to changes in assessment weights (see SI 2.3). Moreover, over 50% of the top 100 compounds underwent a rank change of ≤ 20 positions across all scenarios, demonstrating the robustness of the scoring system (Table 2). Among the scenarios, Scenario 5—focused on occurrence—resulted in the biggest mean absolute and mean relative deviation, in addition to the lowest proportion of compounds with rank changes ≤ 20 positions. This suggests that modifying the scoring related to occurrence had the greatest impact on the final score and ranking of the baseline top 100 compounds, highlighting occurrence as the most influential factor in the CONTRAST prioritisation tool.

Effectiveness of the CONTRAST prioritisation tool for selecting relevant CECs

The effectiveness of the CONTRAST prioritisation tool was evaluated using a list of compounds for validation (335 compounds, see section *Validation of the CONTRAST prioritisation tool* for details). Altogether, 224 of the 268 positive controls (83.6%) were selected using at least one of the schemes when the first approach was applied (Table 3). Most of the positive controls were selected using the T scheme (81.0%). Among these selected compounds, some of them were selected by more than one scheme. From the combination of two or all schemes together, the combination of PB and T schemes resulted in the highest number of positive controls (38.8%). Only four negative controls (11.8%) were selected by either the PM scheme (1 compound) or the T scheme (3 compounds). Based on their final score, they were ranked 2241th or lower. Regarding the unknowns, 10 compounds (29%) were considered to be potentially harmful to the marine

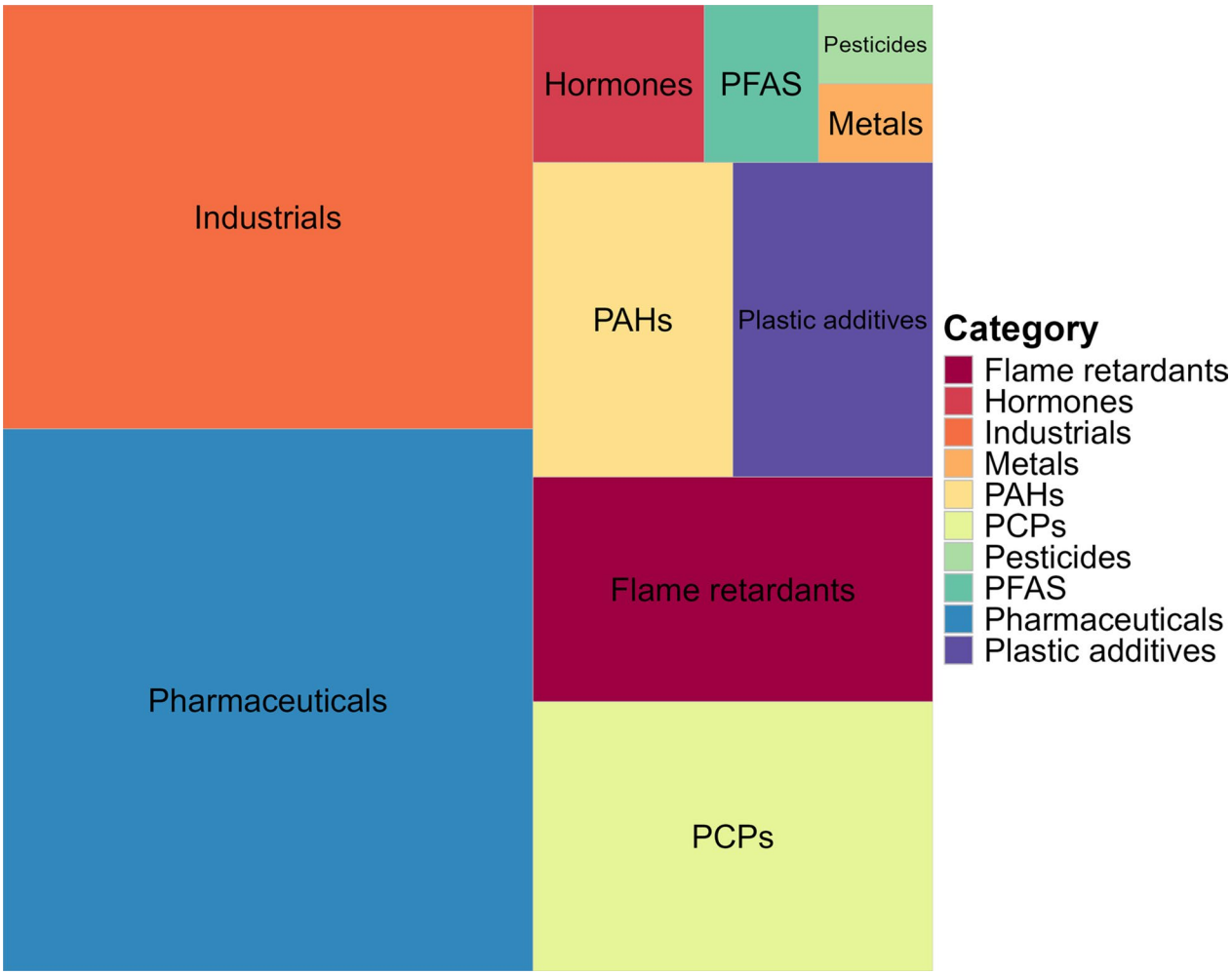

Fig. 5 Diversity of use categories of the top 100 priority compounds. The area of each rectangle is proportional to the number of compounds

environment by the present prioritisation tool. Two unknowns were ranked in the top 100 regulated priority list. Those compounds were diphenylamine and trinitrotoluene, ranked 44th and 93rd respectively.

As for the second approach, overall, less than a quarter of the positive controls were selected (63 compounds, 23.5%). Similar to the first approach, they were mostly filtered by the T scheme (62 compounds, 23.1%). There was only one negative control selected, which was urea, ranked 1524th in the unregulated priority list. Three of the unknowns were also selected, including HMX, nitrogen mustard, and nitroguanidine, ranked 419th, 97th, and 66th respectively. The complete list of selected and not selected compounds and their ranks can be found in SI 2.4.

Discussion

Criteria selected for the CONTRAST prioritisation tool

It is important to develop and apply a marine-specific chemical prioritisation scheme as the differences in environmental conditions can generate different results in the prioritisation depending on how those conditions affect some important characteristic. For instance, the "salting-out" effect in seawater, due to higher salinity compared to freshwater, can increase the bioavailability of certain compounds, thereby elevating their potential to cause harmful effects. As a result, these substances may justify higher prioritisation in marine settings, even though they might pose less risk in freshwater environments. The CECs in the marine environment can also come from marine-specific sources such as shipping, mariculture and antifouling paints [6], further

underscoring the need for a dedicated prioritisation framework.

The survey results showed that toxicity and occurrence were considered as two of the most important features for a CEC to be prioritised because they are directly linked to the potential risk posed to organisms. Mechanisms of toxicity, persistence, and bioaccumulation were regarded as essential to assess the potential impact of CEC. Consideration was also given to the fact that there are environmental dependent schemes based on the source of contaminants, such as prioritisation of CECs in groundwater, drinking water, or in the marine environment. Environmental fate could be used in an exposure-based prioritisation scheme, which considers input, degradation and distribution between media [83]. In the CONTRAST prioritisation, environmental fate was considered linked to persistence, bioaccumulation, and mobility properties assessed during the filtering of compounds.

As bioaccumulation and mobility characteristics of a compound reflect potential accumulation in biota and transport through aquatic environments respectively, combining each of those characteristics with persistence would increase the potential of long-lasting contamination in biota and environment. For this reason, bioaccumulation and mobility were combined with persistence, resulting in PB (Fig. 2) and PM (Fig. 4) schemes. On the other hand, toxic effects can occur due to a single exposure to a compound with high acute toxicity but not necessarily considerable persistence, mobility, and/or bioaccumulation characteristics. Thus, toxicity was assessed using a separate scheme (T scheme, Fig. 3).

While acute and chronic toxicity values are commonly used to assess compound toxicity (SI 1.10 Fig. 2), they don't capture all potential adverse effects on marine organisms. To address this, the CONTRAST prioritisation scheme also incorporated mode of action (MoA), allowing the inclusion of compounds with specific effects like developmental toxicity, mutagenicity, or endocrine disruption—even if their acute/chronic toxicity is low. MoA scoring also distinguishes between compounds with single vs. multiple MoAs, and confirmed vs. suspected MoAs, helping reduce ties in final scores and improving prioritisation, especially with large compound lists.

Based on survey results, certain criteria were excluded from this study. Controlled/regulated compounds like PAHs, PCBs, PBDEs, some pesticides, and PFAS were omitted as they are already well-studied and monitored, with known environmental impacts and regulatory actions in place. The CONTRAST prioritisation tool instead focused on non-regulated compounds to address the knowledge gaps about their distribution, fate, and effects in the marine environment. While both tonnage and emission indicate exposure potential, emission was preferred as it accounts for use patterns and environmental releases, making it more informative. Human health effects were excluded, as they are less relevant for marine-focused assessments.

Outcomes and validation of the CONTRAST prioritisation tool

The results of the application of the CONTRAST prioritisation tool to unregulated substances in the EU showed that 6PPD was the highest-ranked compound. This compound is widely used as an antioxidant and antiozonant in tyres. 6PPD and its derivative 6PPD-quinone are gaining more attention due to their presence in the environment and potential adverse effects [84]. 6PPD was selected by PB and T schemes owing to its potential bioaccumulation and persistence, high (eco)toxicity, and multiple modes of action (developmental toxicity and mutagenicity). While having higher stability than 6PPD [85], 6PPD-quinone was not present in the list of unregulated priority compounds because this compound was not yet included in the version of PikMe used in this study.

Despite the exclusion of regulated compounds, some of the legacy contaminants were not filtered out and ranked high in the CONTRAST priority list, including PAHs, PBDEs, alkanes, and PFAS. Those legacy contaminants are regularly monitored but not yet added to the list of hazardous substances in EU regulations. For instance, fluorene and perylene, both PAHs, were ranked 2nd and 3rd respectively, due to their high scores in persistence and bioaccumulation, toxicity, and occurrence. Those compounds are listed as active substances in the US EPA Toxic Substances Control Act (TSCA) [86]. However, they were not selected as indicator compounds for environmental monitoring. Thus, they are not listed as priority substances in EU regulations and were not filtered out in the CONTRAST prioritisation schemes.

Regarding the diversity of use categories for the compounds that ranked high, pharmaceuticals constituted a major proportion of compounds in the top 100 unregulated priority compounds. Although pharmaceuticals are active substances that are intended to treat disease or illness in humans or animals, they can also provoke potentially harmful changes in non-target organisms when they enter the environment [87]. The potential toxic effects of some pharmaceuticals explain their high scores on toxicity and/or MoA in the CONTRAST prioritisation. In addition, pharmaceuticals were also one of the most frequently reported group of chemicals in the NORMAN EMPODAT and Norwegian Monitoring databases, alongside industrial chemicals and biocides. However, many of the latter compounds were filtered out

from the top 100 priority compounds due to their regulated status.

The clustering based on the scores from each assessment showed that the top 100 unregulated and regulated priority lists consisted of several clusters characterised by dominant P, B, T, and M properties, MoA, occurrence, or emission. The diversity in criteria that influenced the clustering implied that the CONTRAST prioritisation was not biased by only one criterion and that the scoring system was effective in giving equal weight to the criteria used for assessing the compounds. Almost half of the compounds in the top 100 unregulated priority list are persistent, bioaccumulative, and toxic.

More than half of the top 100 compounds demonstrated robustness to changes in scoring, underscoring the reliability of the CONTRAST prioritisation tool. The sensitivity analysis identified the occurrence assessment as the most influential factor affecting the prioritisation outcome. Notably, rank shifts of ≥ 20 positions were observed in certain compound groups—such as PBDEs, PFAS, and organophosphate esters—which received the highest occurrence scores (i.e. 2, presence > LOQ in both the NORMAN EMPODAT and Norwegian monitoring databases). As their occurrence scores contributed significantly to their overall prioritisation scores, adjustments to the occurrence weighting had a substantial impact on their rankings.

Validation of the PB, T, and PM schemes showed that over 80% of positive controls (i.e. compounds prioritised under European regulations and RSCs) were successfully selected when regulated compounds were included (Table 3), confirming the schemes' effectiveness. However, since CONTRAST focuses on identifying unregulated CECs, regulated compounds were excluded before applying the schemes. This reduced the selection of positive controls to under 25%, demonstrating that the exclusion effectively narrows the focus to less-studied compounds. Most selected positive controls were listed under RSCs or the WFD Watch Lists, which target emerging pollutants with limited monitoring data.

Four negative controls (i.e. compounds with low environmental risk) were selected when regulated compounds were included; only one—urea—was selected when they were excluded. Urea was selected by the PM scheme due to its persistence and mobility, despite being non-toxic and non-bioaccumulative. Its high environmental release stems from widespread use as a fertilizer. Additionally, three "unknowns" (regulated compounds not necessarily harmful) were selected by the T scheme under both inclusion and exclusion approaches. Among these, nitroguanidine ranked highest due to its selection by both the T and PM schemes, indicating toxicity along with persistence and mobility.

A comparison between the number of compounds in the existing priority lists that were selected by one of the three schemes showed that more compounds were selected by the T scheme compared with the PB and PM schemes, which is explained by the use of two approaches based on ecotoxicity (acute and chronic toxicities) as well as MoA in the T scheme. Those two approaches were applied to cover a broader range of potential adverse effects to marine organisms. As only a small number of compounds have been tested using non-standard ecotoxicity endpoints, filtering based on MoA relied on predicted data, which are available for a large number of compounds in the PikMe database.

Strengths and limitations of the CONTRAST prioritisation tool

The CONTRAST prioritisation tool had several strengths for application to CECs in the marine environment compared with the other prioritisation tools reviewed in this study. First, the CONTRAST prioritisation tool included a step to exclude regulated compounds to focus more on non-regulated compounds that are less studied, thus having less available data. This step ensured that the output of the CONTRAST prioritisation consisted of compounds with knowledge gaps regarding their distribution, fate, and effects, highlighting the need for further research. Experimental work on these compounds within the CONTRAST project may lead to justification for their inclusion in regular monitoring programmes or as hazardous compounds in the European regulation.

Second, MoA had a considerable weight in the CONTRAST prioritisation tool as it was part of the filtering using the T scheme and had a separate assessment and scoring for the ranking of selected compounds. Most of the existing prioritisation schemes identified in the literature review used acute and chronic toxicity values to assess the toxicity of a compound. However, the CONTRAST prioritisation tool went a step further by also including non-standard endpoints, such as endocrine disrupting potential, carcinogenicity, mutagenicity, reproductive toxicity, specific target organ toxicity, and developmental toxicity, from experimental and predicted data.

Unlike other prioritisation schemes that begin with limited compound lists from monitoring or suspect databases, the CONTRAST prioritisation tool started with a broad initial list of 1.13 million chemicals from the PikMe database [88]. This expanded scope enables the selection of widely detected parent compounds, potentially harmful transformation products, and unmonitored substances that may threaten marine ecosystems. The comprehensive coverage makes the CONTRAST prioritisation tool applicable across different sea regions,

though region-specific occurrence data may be needed for non-European areas.

Another strength is that the CONTRAST prioritisation tool gave equal weight to both regularly monitored (after filtering out regulated substances) and less studied compounds. This was done by assessing occurrence based on the absence or presence in two environmental monitoring databases instead of the number of measurements used by many prioritisation schemes. As CECs generally have not been widely studied, such substances often have relatively little or no occurrence data in comparison to regularly monitored contaminants. The assessment based on absence or presence in environmental monitoring databases ensured that less-studied CECs weren't ranked lower due to missing occurrence data.

Despite these strengths, the output of the CONTRAST prioritisation is not the ultimate list that can be directly used for monitoring or in experimental tests. Due to some considerations in analytical and toxicological aspects (i.e. some compounds are unstable in environmental matrices or not relevant for experimental studies using certain organisms due to limited absorption), not all the compounds in the CONTRAST priority list were relevant for further studies. In addition, some legacy contaminants were still present in the top 100 unregulated priority compounds, regardless of the exclusion of regulated compounds. Thus, expert judgments were needed to decide which traditional compounds should be excluded from the priority list due to the abundance of available data on their occurrence, fate, and effects.

The CONTRAST prioritisation is a dynamic tool that depends on input data regarding P, B, M, and T properties, as well as occurrence, MoA, and emission. Therefore, future results obtained using the tool will be affected by updates on the databases, the progress in knowledge with respect to input data, and changes in legislation. For example, many PFAS had low or zero PB and/or PM scores despite numerous studies proving their extreme persistence. Due to the lack of data on the persistence of PFAS in the version of PikMe used in this study, those compounds were not selected by the PB and/or PM scheme, resulting in a score of zero. This issue can be addressed by applying the CONTRAST prioritisation tool to the updated version of the PikMe database. As the current study used an initial list of 1.13 million compounds for which data were mostly generated using predictions, more experimental data in the future would change the outcome of the tool.

The same reasoning can be applied to the evolution of legislation, which could change the outcome of the CONTRAST prioritisation. Some compounds that are currently not regulated may be considered for inclusion in future regulatory hazard lists. A proposal amending WFD and EQS (published in October 2022) includes additional substances, such as 17-Beta estradiol, clarithromycin, and ibuprofen [89], which are present in the top 100 CONTRAST priority list. As this proposal might be approved in the future, those compound can become regulated and, in this case, can no longer be considered as CECs according to this study. Conversely, emergency authorisation can be granted to some banned compounds, as previously happened with neonicotinoid biocides, such as imidacloprid, thiamethoxam, and clothianidin [90].

One of the limitations of the CONTRAST prioritisation tool is the lack of reliable data for emission and occurrence assessments. Tonnage estimates are often outdated, and over 86% of compounds have an "unknown" functional category in the ChemExpo Knowledgebase, resulting in low use pattern scores. The absence of occurrence data of more than 10% of the CONTRAST priority compounds and seasonal variations in marine environments need to be addressed. These gaps are expected to narrow as databases are updated and more data becomes available. Additionally, no distinction in scoring was made between different sources of measured and estimated data for the MoA assessment despite different method and predictive models employed. Therefore, regular updates to the CONTRAST prioritisation using additional certainty scores in MoA and current data are essential to reflect evolving knowledge and regulations.

Conclusion

The marine environment has different environmental conditions, notably salinity and pH, affecting partitioning, sorption, and bioavailability of compounds. Due to the lack of prioritisation schemes that consider these specificities of the marine environment and their effects on the bioaccumulation of compounds, a marine-specific prioritisation tool was developed in this study using filtering with adapted bioaccumulation criteria followed by scoring and ranking approaches. The filtering process was carried out using three parallel schemes (PB, T, and PM schemes), and the scoring was based on MoA, occurrence, and emission.

As a marine-specific prioritisation tool, the CONTRAST prioritisation takes into account marine sources as a specific type of chemical emissions. Furthermore, the CONTRAST prioritisation tool covers a broader list of possible chemicals that can be present in the marine environment (1.13 million chemicals in the PikMe tool database) and adverse effects of CECs on organisms by integrating a wide range of modes of action besides standard ecotoxicity-based assessment. The potential for future research to fill the knowledge gaps was taken into consideration by giving an equal value to compounds

with limited occurrence data and filtering out the EU-regulated compound whose toxicity is well-known.

The CONTRAST prioritisation tool was able to select and prioritise 8548 unregulated priority compounds from the 1.13 million compounds in the PikMe database, ranking them based on their final score. The top 100 unregulated priority list were dominated by pharmaceuticals and industrial chemicals but also included other use categories, such as flame retardants, plasticisers, PCPs, or biocides. The diversity of compounds in the top 100 unregulated priority list concerned not only the use categories but also the criteria that influence their final score, indicating that the CONTRAST prioritisation tool was not biased by any single criterion. The validation of the filtering approach revealed that more than half of compounds from several priority lists in European legislations and RSCs were selected if the regulated compounds were not excluded, demonstrating the effectiveness of the tool in selecting hazardous compounds.

The CONTRAST prioritisation served its purpose as a tool to select CompTox Chemical Dashboard CECs for the development of an integrated assessment framework in the CONTRAST project. Moreover, it can support regulation by guiding the decision-making process regarding which compounds should be monitored in the marine environment and studied for their fate, distribution, and effects. The CONTRAST prioritisation tool was developed with a focus on Europe, for applications within the Marine Strategy Framework Directive (MSFD) or European RSCs. However, it can also be applied on a broader scale within other RSCs by adapting the input data, e.g. using regional-specific occurrence and emission data as well as adapting the list of excluded compounds based on current regulations in other regions.

Disclaimer

This study and its conclusion represent the opinions of the authors, but not necessarily those of the organisations they work for, or an endorsement of the tools and methods used therein.

Abbreviations

ACE	Assessment of Antifouling Agents in Coastal Environments
AD	Applicability domain
ADI	Acceptable daily intake
AMAP	Artic Monitoring and Assessment
BCF	Bioconcentration factor
BPR	Biocidal Product Regulation
CAESAR	Computer Assisted Evaluation of industrial chemical Substances According to Regulations
CAS	Chemicals Abstracts Service
CECs	Contaminants of emerging concern
CERAPP	Collaborative Estrogen Receptor Activity Prediction Project
ChV	Chronic value
CLP	Classification, labelling, and packaging
CoMPARA	Collaborative Modelling Project for Androgen Receptor Activity
CompTox	Computational toxicology
CONTRAST	Contaminants of Emerging Concern: An Integrated Approach for Assessing Impacts on the Marine Environment
CoRAP	Community rolling action
EC_{10}	10% Effective concentration
EC_{50}	Half maximal effective concentration
ECHA	European Chemical Agency
ECs	Emerging contaminants
EEA	European Economic Area
EMPODAT	Database of geo-referenced monitoring and bio-monitoring data on emerging substances in air, water and soil
EMSA	European Maritime Safety Agency
EPA	Environmental Protection Agency
EPI Suite	Estimation Programs Interface Suite
ERC	Environmental release category
EU	European Union
GMM	Gaussian mixture model
HELCOM	Helsinki Commission
K_{OA}	Octanol–air partition coefficient
K_{OC}	Organic carbon/water partition coefficient
K_{OW}	N-Octanol/water partition coefficient
LC_{50}	Lethal concentration for 50% of the tested group
LD_{50}	Lethal dose for 50% of the tested group
LOD	Limit of detection
LOEC	Lowest observed effect level
LOQ	Limit of quantification
M	Mobility
MEC	Measured environmental concentration
MRL	Minimal risk level
MSFD	Marine Strategy Framework Directive
NIVA	Norwegian Institute for Water Research
NOEC	No observed effect concentration
NORMAN	Network of reference laboratories, research centres and related organisations for monitoring of emerging environmental substances
OC	Organic Carbon
OECD TG	Organization for Economic and Co-operation and Development technical guidance
OPERA	Open (Quantitative) Structure–activity/property Relationship App
OSPAR	Convention for the Protection of the Marine Environment of the North-East Atlantic
PAHs	Polycyclic aromatic hydrocarbons
PB	Persistent and bioaccumulative
PBDEs	Polybrominated diphenyl ethers
PBT	Persistent, bioaccumulative, and toxic
PC	Principal component
PCA	Principal component analysis
PCBs	Polychlorinated biphenyls
PCPs	Personal care products
PEC	Predicted environmental concentration
PFAS	Per- and polyfluoroalkyl substances
PikMe	Project called "Prioritization, identification, and quantification of new environmental contaminants efficiently"
PIC	Prior Informed Consent Regulation
PM	Persistent and mobile
PMOCs	Persistent and mobile organic compounds
PNEC	Predicted no effect concentration
$PNEC_{hum}$	Predicted no effect concentration on human health
POPs	Persistent organic pollutants
potB	Potentially bioaccumulative
potI	Potentially impactful
potP	Potentially persistent
QSAR	Quantitative structure–activity relationship
REACH	Registration, Evaluation, Authorization, and Restriction of Chemicals
RfD	Reference dose
RSCs	Regional Sea Conventions
S(i)	Silhouette index
SI	Supplementary information
SPIN	Substances in Preparation in Nordic Countries
SVHC	Substances of very high concern

T	Toxicity
TEST	Toxicity Estimation Software Tool
TQs	Toxicity quotients
TSCA	Toxic Substances Control Act
t-SNE	T-Distributed stochastic neighbour embedding
UMAP	Uniform manifold approximation and projection
vB	Very bioaccumulative
vM	Very mobile
vP	Very persistent
vPvB	Very persistent and very bioaccumulative
WFD	Water Framework Directive
WCSS	Within-cluster sum of squares
WWTP	Wastewater treatment plants

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12302-025-01257-9.

Additional file 1.

Additional file 2.

Acknowledgements
The authors would like to acknowledge The Norwegian Institute for Water Research (NIVA) and The Climate and Environmental Research Institute NILU for giving permission to use the PikMe database during this study.

Author contributions
PYY, BDW, AA, JB, JB, NB, SB, ASB, MAF, KH, HK, VML, SM, AM, FM, MR, JS, SV, ACW conceptualise the CONTRAST prioritisation tool. PYY did the data analysis. PYY wrote the initial manuscript. BDW contributed to and improved the manuscript. BDW, AA, JB, JB, NB, SB, ASB, MAF, KH, HK, VML, SM, AM, FM, MR, JS, CVP, SV, CW, KD, ACW read and revised the manuscript.

Funding
The CONTRAST project is funded by the European Union's Horizon-CL6-2023-ZEROPOLLUTION-01(Clean environment and zero pollution) grant agreement No. 101135037-CONTRAST-Horizon CL6-2023.

Availability of data and materials
The PikMe database analysed during the current study was a preliminary version of the PikMe database and tool which will be publicly available on Zenodo (https://doi.org/10.5281/zenodo.15647470) [88]. The data can, however, currently be accessed upon reasonable request and with the permission of the PikMe database's authors. The development of this tool was funded by the Norwegian Environmental Agency.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Marine Research (ILVO-Marine), Jacobsenstraat 1, 8400 Ostend, Belgium. [2]Institute of Marine Research (IMR), Contaminants and Biohazards, Nordnesgaten 50, 5817 Bergen, Norway. [3]Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft NR330HT, UK. [4]Centro Oceanográfico de Vigo, Instituto Español de Oceanografía (IEO-CSIC), Subida a Radio Faro 50, 36390 Vigo, Spain. [5]French National Institute for Ocean Science and Technology (IFREMER), Chemical Contamination of Marine Ecosystem, Rue de L'Ile d'Yeu, BP 21105, 44311 Nantes, France. [6]Ecotoxicology and Risk Assessment, Norwegian Institute for Water Research (NIVA), Økernveien 94, 0579 Oslo, Norway. [7]National Institute of Oceanography and Applied Geophysics IT (OGS), Borgo Grotta Gigante 42/C, 34010 Sgonico-Trieste, Italy. [8]Institute of Oceanography (HCMR), Hellenic Centre for Marine Research, Leoforos Athens Sounio 467 Km, 19013 Attikia Anavissos, Greece. [9]Centro Oceanográfico de Murcia, Instituto Español de Oceanografía (IEO-CSIC), C/ Varadero 1, 30740 Murcia, Spain. [10]Department of Biological and Environmental Sciences, University of Gothenburg, Box 463, 40530 Gothenburg, Sweden. [11]Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Technology and Food Science, Brusselsesteenweg, 370, 9090 Merelbeke-Melle, Belgium. [12]Department of Green Chemistry and Technology, Research Group Environmental Organic Chemistry and Technology (EnVOC), Ghent University, Coupure Links 653, 9000 Ghent, Belgium.

Received: 15 August 2025 Accepted: 25 October 2025
Published online: 18 November 2025

References

1. Chemical Abstracts Service (2025) CAS Registry[SM] Homepage. https://www.cas.org/cas-data/cas-registry. Accessed 3 April 2025.
2. Ruan T, Li PY, Wang HT, Li TY, Jiang GB (2023) Identification and prioritization of environmental organic pollutants: from an analytical and toxicological perspective. Chem Rev 123(17):10584–10640. https://doi.org/10.1021/acs.chemrev.3c00056
3. ECHA (2024) REACH registration statistics. https://echa.europa.eu/registration-statistics. Accessed 18 June 2024.
4. Nipen M, Vogt RD, Bohlin-Nizzetto P, Borgå K, Mwakalapa EB, Borgen AR, Schlabach M, Christensen G, Mmochi AJ, Breivik K (2022) Increasing trends of legacy and emerging organic contaminants in a dated sediment core from East-Africa. Front Environ Sci. https://doi.org/10.3389/fenvs.2021.805544
5. Tong X, Mohapatra S, Zhang J, Tran NH, You L, He Y, Gin KY (2022) Source, fate, transport and modelling of selected emerging contaminants in the aquatic environment: current status and future perspectives. Water Res 217:118418. https://doi.org/10.1016/j.watres.2022.118418
6. Tornero V, Hanke G (2016) Chemical contaminants entering the marine environment from sea-based sources: a review with a focus on European seas. Mar Pollut Bull 112(1–2):17–38. https://doi.org/10.1016/j.marpolbul.2016.06.091
7. León VM, Bellas J (2023) Introduction. In: León VM, Bellas J (eds) Contaminants of emerging concern in the marine environment. Elsevier, pp 1–10. https://doi.org/10.1016/B978-0-323-90297-7.00011-1
8. James CA, Sofield R, Faber M, Wark D, Simmons A, Harding L, O'Neill S (2023) The screening and prioritization of contaminants of emerging concern in the marine environment based on multiple biological response measures. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2023.163712
9. Barbosa J, De Schamphelaere K, Janssen C, Asselman J (2021) Prioritization of contaminants and biological process targets in the North Sea using toxicity data from ToxCast. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2020.144157
10. Chiaia-Hernández AC, Scheringer M, Müller A, Stieger G, Wächter D, Keller A, Pintado-Herrera MG, Lara-Martin PA, Bucheli TD, Hollender J (2020) Target and suspect screening analysis reveals persistent emerging organic contaminants in soils and sediments. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2020.140181
11. Park N, Jeon J (2021) Emerging pharmaceuticals and industrial chemicals in Nakdong River, Korea: identification, quantitative monitoring, and prioritization. Chemosphere. https://doi.org/10.1016/j.chemosphere.2020.128014
12. Plaza-Hernández M, Legler J, MacLeod M (2021) Integration of production and use information into an exposure-based screening approach to rank chemicals of emerging Arctic concern for potential to be planetary boundary threats. Emerg Contam 7:213–218. https://doi.org/10.1016/j.emcon.2021.10.001
13. Schulze S, Sättler D, Neumann M, Arp HPH, Reemtsma T, Berger U (2018) Using REACH registration data to rank the environmental emission potential of persistent and mobile organic chemicals. Sci Total Environ 625:1122–1128. https://doi.org/10.1016/j.scitotenv.2017.12.305

14. Montes R, Méndez S, Carro N, Cobas J, Alves N, Neuparth T, Santos MM, Quintana JB, Rodil R (2022) Screening of contaminants of emerging concern in surface water and wastewater effluents, assisted by the persistency-mobility-toxicity criteria. Molecules. https://doi.org/10.3390/molecules27123915
15. Sanderson H, Fauser P, Bengström L, Vorkamp K (2024) Semi-quantitative risk-based prioritisation scheme for chemicals of concern in the Nordic countries. RSC Sustain 2:558–566. https://doi.org/10.1039/d3su00322a
16. Zhou J, Zuo SQ, Wang Y, An RQ, Zheng SP, Cao LZ, Zhang F, Jin YH, Shen J, Meng HY, Zheng XH, Wang CC, Shi HC, Zhao YB, Giesy JP, Zhang K, Dai JY (2023) Prioritization of risks posed by synthetic chemicals manufactured in China toward humans and the environment via persistence, bioaccumulation, mobility and toxicity properties. Environ Int. https://doi.org/10.1016/j.envint.2023.108042
17. Utami RR, Geerling GW, Salami IRS, Notodarmojo S, Ragas AMJ (2020) Environmental prioritization of pesticide in the Upper Citarum River Basin, Indonesia, using predicted and measured concentrations. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2020.140130
18. Tsaboula A, Papadakis EN, Vryzas Z, Kotopoulou A, Kintzikoglou K, Papadopoulou-Mourkidou E (2016) Environmental and human risk hierarchy of pesticides: a prioritization method, based on monitoring, hazard assessment and environmental fate. Environ Int 91:78–93. https://doi.org/10.1016/j.envint.2016.02.008
19. Wen Y, Huang SY, Qin ZH, Chen ZL, Shao Y (2024) Priority screening on emerging contaminants in sediments of the Yangtze River, China. Environ Sci Eur. https://doi.org/10.1186/s12302-024-00855-3
20. Daginnus K, Gottardo S, Payá-Pérez A, Whitehouse P, Wilkinson H, Zaldívar JM (2011) A model-based prioritisation exercise for the European Water Framework Directive. Int J Environ Res Public Health 8(2):435–455. https://doi.org/10.3390/ijerph8020435
21. Dulio V, Alygizakis N, Ng K, Schymanski EL, Andres S, Vorkamp K, Hollender J, Finckh S, Aalizadeh R, Ahrens L, Bouhoulle E, Cirka L, Derksen A, Deviller G, Duffek A, Esperanza M, Fischer S, Fu QG, Gago-Ferrero P, Haglund P, Junghans M, Kools SAE, Koschorreck J, Lopez B, de Alda ML, Mascolo G, Miège C, Osté L, O'Toole S, Rostkowski P, Schulze T, Sims K, Six L, Slobodnik J, Staub PF, Stroomberg G, Thomaidis NS, Togola A, Tomasi G, von der Ohe PL (2024) Beyond target chemicals: updating the NORMAN prioritisation scheme to support the EU chemicals strategy with semi-quantitative suspect/non-target screening data. Environ Sci Eur. https://doi.org/10.1186/s12302-024-00936-3
22. Billings A, Jones KC, Pereira MG, Spurgeon DJ (2024) Emerging and legacy plasticisers in coastal and estuarine environments: a review. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2023.168462
23. Romain SJS, Basirico LM, Hutton S, Brander S, Armbrust KL (2024) Influence of salinity on the partitioning behavior of six commonly used pesticides in fish eggs. Environ Toxicol Chem 43(2):299–306. https://doi.org/10.1002/etc.5780
24. Saranjampour P, Vebrosky EN, Armbrust KL (2017) Salinity impacts on water solubility and N-octanol/water partition coefficients of selected pesticides and oil constituents. Environ Toxicol Chem 36(9):2274–2280. https://doi.org/10.1002/etc.3784
25. Xie W-H, Shiu W-Y, Mackay D (1997) A review of the effect of salts on the solubility of organic compounds in seawater. Mar Environ Res 44(4):429–444. https://doi.org/10.1016/S0141-1136(97)00017-2
26. Bayen S, Zhang H, Desai MM, Ooi SK, Kelly BC (2013) Occurrence and distribution of pharmaceutically active and endocrine disrupting compounds in Singapore's marine environment: influence of hydrodynamics and physical-chemical properties. Environ Pollut 182:1–8. https://doi.org/10.1016/j.envpol.2013.06.028
27. Marion GM, Millero FJ, Camões MF, Spitzer P, Feistel R, Chen CTA (2011) pH of seawater. Mar Chem 126(1–4):89–96. https://doi.org/10.1016/j.marchem.2011.04.002
28. US EPA (2024) pH. https://www.epa.gov/caddis/ph. Accessed 29 July 2024.
29. Ojemaye CY, Petrik L (2019) Pharmaceuticals in the marine environment: a review. Environ Rev 27(2):151–165. https://doi.org/10.1139/er-2018-0054
30. Nybom I, Horlitz G, Gilbert D, Berrojalbiz N, Martens J, Arp HPH, Sobek A (2021) Effects of organic carbon origin on hydrophobic organic contaminant fate in the Baltic Sea. Environ Sci Technol 55(19):13061–13071. https://doi.org/10.1021/acs.est.1c04601
31. Yulikayani PY, De Witte B, Ali A, Barber J, Bellas J, Briant N, Brooks S, Bruvold AS, French MA, Hylland K, Kaberi H, León VM, Martins S, Maufret A, Molinari F, Reid MJ, Sturve J, Valdersnes S, Wennberg AC (2025) Prioritisation of CECs in the marine environment using hazard-based filtering and ranking approaches. In: SETAC Europe 35th Annual Meeting, Vienna, Austria, 11–15 May 2025.
32. Bryer J, Speerschneider K (2016) likert: analysis and visualization likert items_. R package version 1.3.5. https://CRAN.R-project.org/package=likert.
33. R Core Team (2023) R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.
34. Wennberg AC, Rostkowski P, Reid MJ (2025) PikMe: a flexible prioritization tool for chemicals of emerging concern. Environ Toxicol Chem (Accepted). https://doi.org/10.1093/etojnl/vgaf237
35. European Commission (2024) Active substances, safeners and synergists. https://ec.europa.eu/food/plant/pesticides/eu-pesticides-database/start/screen/active-substances. Accessed 16 September 2024.
36. ECHA (2023) Guidance on information requirements and chemical safety assessment chapter R.11: PBT/vPvB assessment. Version 4.0.
37. European Commission (2023) Commission Delegated Regulation (EU) 2023/707 of 19 December 2022 amending Regulation (EC) No 1272/2008 as regards hazard classes and criteria for the classification, labelling and packaging of substances and mixtures. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32023R0707&from=EN. Accessed 2 May 2024.
38. European Commission (2006) REGULATION (EC) No 1907/2006 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. https://eur-lex.europa.eu/eli/reg/2006/1907/oj/eng.
39. Martin T (2016) Toxicity Estimation Software Tool (TEST). U.S. Environmental Protection Agency. Washington, DC. https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test.
40. Cassano A, Manganaro A, Martin T, Young D, Piclin N, Pintore M, Bigoni D, Benfenati E (2010) CAESAR models for developmental toxicity. Chem Cent J. https://doi.org/10.1186/1752-153x-4-S1-S4
41. Hansen K, Mika S, Schroeter T, Sutter A, ter Laak A, Steger-Hartmann T, Heinrich N, Müller KR (2009) Benchmark data set for in silico prediction of Ames mutagenicity. J Chem Inf Model 49(9):2077–2081. https://doi.org/10.1021/ci900161g
42. Mansouri K, Abdelaziz A, Rybacka A, Roncaglioni A, Tropsha A, Varnek A, Zakharov A, Worth A, Richard AM, Grulke CM, Trisciuzzi D, Fourches D, Horvath D, Benfenati E, Muratov E, Wedebye EB, Grisoni F, Mangiatordi GF, Incisivo GM, Hong HX, Ng HW, Tetko IV, Balabin I, Kancherla J, Shen J, Burton J, Nicklaus M, Cassotti M, Nikolov NG, Nicolotti O, Andersson PL, Zang QD, Politi R, Beger RD, Todeschini R, Huang RL, Farag S, Rosenberg SA, Slavov S, Hu X, Judson RS (2016) CERAPP: collaborative estrogen receptor activity prediction project. Environ Health Persp 124(7):1023–1033. https://doi.org/10.1289/ehp.1510267
43. Mansouri K, Kleinstreuer N, Abdelaziz AM, Alberga D, Alves VM, Andersson PL, Andrade CH, Bai F, Balabin I, Ballabio D, Benfenati E, Bhhatarai B, Boyer S, Chen J, Consonni V, Farag S, Fourches D, García-Sosa AT, Gramatica P, Grisoni F, Grulke CM, Hong H, Horvath D, Hu X, Huang R, Jeliazkova N, Li J, Li X, Liu H, Manganelli S, Mangiatordi GF, Maran U, Marcou G, Martin T, Muratov E, Nguyen D-T, Nicolotti O, Nikolov NG, Norinder U, Papa E, Petitjean M, Piir G, Pogodin P, Poroikov V, Qiao X, Richard AM, Roncaglioni A, Ruiz P, Rupakheti C, Sakkiah S, Sangion A, Schramm K-W, Selvaraj C, Shah I, Sild S, Sun L, Taboureau O, Tang Y, Tetko IV, Todeschini R, Tong W, Trisciuzzi D, Tropsha A, Driessche GVD, Varnek A, Wang Z, Wedebye EB, Williams AJ, Xie H, Zakharov AV, Zheng Z, Judson RS (2020) CoMPARA: collaborative modeling project for androgen receptor activity. Environ Health Persp 128(2):027002. https://doi.org/10.1289/EHP5580
44. Norman Network NORMAN EMPODAT Database-Chemical Occurrence Data. https://www.norman-network.com/nds/empodat/. Accessed 21 May 2024.

45. Dimitrov S, Dimitrova G, Pavlov T, Dimitrova N, Patlewicz G, Niemela J, Mekenyan O (2005) A stepwise approach for defining the applicability domain of SAR and QSAR models. J Chem Inf Model 45(4):839–849. https://doi.org/10.1021/ci0500381
46. Mansouri K, Grulke CM, Judson RS, Williams AJ (2018) Opera models for predicting physicochemical properties and environmental fate endpoints. J Cheminform 10(1):10. https://doi.org/10.1186/s13321-018-0263-1
47. Norwegian Environmental Agency (2024) Nowegian monitoring database Vannmiljø. https://vannmiljo.miljodirektoratet.no/. Accessed July 2024.
48. Dulio V, von der Ohe PC (2013) NORMAN Prioritisation framework for emerging substances. NORMAN Association, Verneuil-en-Halatte, France
49. US EPA (2024) Function Categories. ChemExpo Knowledgebase. 1.0. https://comptox.epa.gov/chemexpo/. Accessed 23 Aug 2024.
50. Dionisio KL, Phillips K, Price PS, Grulke CM, Williams A, Biryol D, Hong T, Isaacs KK (2018) The chemical and products database, a resource for exposure-relevant data on chemicals in consumer products. Sci Data 5(1):180125. https://doi.org/10.1038/sdata.2018.125
51. Tornero V, Hanke G (2016) Identification of marine chemical contaminants released from sea-based sources: A review focusing on regulatory aspects. JRC102452. EUR 28039. Luxembourg, Luxembourg.
52. European Commission (2015) Commission Implementing Decision (EU) 2015/495 of 20 March 2015 establishing a watch list of substances for Union-wide monitoring in the field of water policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council. https://eur-lex.europa.eu/eli/dec_impl/2015/495/oj/eng.
53. European Commission (2018) Commission Implementing Decision (EU) 2018/840 of 5 June 2018 establishing a watch list of substances for Union-wide monitoring in the field of water policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council andrepealing Commission Implementing Decision (EU) 2015/495. https://eur-lex.europa.eu/eli/dec_impl/2018/840/oj/eng.
54. European Commission (2020) Commission Implementing Decision (EU) 2020/1161 of 4 August 2020 establishing a watch list of substances for Union-wide monitoring in the field of water policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council. https://eur-lex.europa.eu/eli/dec_impl/2020/1161/oj/eng.
55. European Commission (2022) Commission Implementing Decision (EU) 2022/1307 of 22 July 2022 establishing a watch list of substances for Union-wide monitoring in the field of water policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council. https://eur-lex.europa.eu/eli/dec_impl/2022/1307/oj/eng.
56. Gomez Cortes L, Porcel Rodriguez E, Marinov D, Sanseverino I, Lettieri T (2025) Selection of substances for the 5th Watch List under the Water Framework Directive. Publication Office of the European Union, Luxembourg. https://doi.org/10.2760/956398
57. OSPAR Commission (2024) OSPAR Agreement 2013-06: OSPAR list of substances used and discharged offshore which are considered to pose little or no risk to the environment (PLONOR) - Update 2024. https://www.ospar.org/documents?v=32939. Accessed 12 June 2025.
58. Osman R, Saim N, Juahir H, Abdullah MP (2012) Chemometric application in identifying sources of organic contaminants in Langat river basin. Environ Monit Assess 184(2):1001–1014. https://doi.org/10.1007/s10661-011-2016-8
59. Jiménez-Preciado AL, Cruz-Aké S, Venegas-Martínez F (2024) Identification of patterns in CO2 emissions among 208 countries: k-means clustering combined with PCA and non-linear t-SNE visualization. Mathematics. https://doi.org/10.3390/math12162591
60. Mehrjou A, Hosseini R, Nadjar Araabi B (2016) Improved Bayesian information criterion for mixture model selection. Pattern Recogn Lett 69:22–27. https://doi.org/10.1016/j.patrec.2015.10.004
61. Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7
62. Reppas-Chrysovitsinos E, Sobek A, MacLeod M (2017) Screening-level exposure-based prioritization to identify potential POPs, vPvBs and planetary boundary threats among Arctic contaminants. Emerg Contam 3:85–94. https://doi.org/10.1016/j.emcon.2017.06.001
63. Johnson AC, Donnachie RL, Sumpter JP, Jurgens MD, Moeckel C, Pereira MG (2017) An alternative approach to risk rank chemicals on the threat they pose to the aquatic environment. Sci Total Environ 599:1372–1381. https://doi.org/10.1016/j.scitotenv.2017.05.039
64. Kuzmanovic M, Ginebreda A, Petrovic M, Barcelo D (2015) Risk assessment based prioritization of 200 organic micropollutants in 4 Iberian rivers. Sci Total Environ 503:289–299. https://doi.org/10.1016/j.scitotenv.2014.06.056
65. OSPAR Commission (2006) Dynamic selection and prioritisation mechanism for hazardous substances (new DYNAMEC manual).
66. Arp HPH, Hale S, Borchers U, Valkov V, Wiegand L, Zahn D, Neuwald I, Nödler K, Scheurer M (2023) A prioritization framework for PMT/vPvM Substances under REACH for registrants, regulators, researchers and the water sector. German Environment Agency. Dessau-Roßlau, Germany.
67. Verlicchi P, Grillini V, Lacasa E, Archer E, Krzeminski P, Gomes AI, Vilar VJP, Rodrigo MA, Gäbler J, Schafer L (2023) Selection of indicator contaminants of emerging concern when reusing reclaimed water for irrigation - a proposed methodology. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2023.162359
68. von der Ohe PC, Dulio V, Slobodnik J, De Deckere E, Kühne R, Ebert RU, Ginebreda A, De Cooman W, Schüürmann G, Brack W (2011) A new risk assessment approach for the prioritization of 500 classical and emerging organic microcontaminants as potential river basin specific pollutants under the European water framework directive. Sci Total Environ 409(11):2064–2077. https://doi.org/10.1016/j.scitotenv.2011.01.054
69. Maloney EM, Villeneuve DL, Blackwell BR, Vitense K, Corsi SR, Pronschinske MA, Jensen KM, Ankley GT (2023) A framework for prioritizing contaminants in retrospective ecological assessments: application in the Milwaukee Estuary (Milwaukee, WI). Integr Environ Assess 19(5):1276–1296. https://doi.org/10.1002/ieam.4725
70. Arp HPH, Brown TN, Berger U, Hale SE (2017) Ranking REACH registered neutral, ionizable and ionic organic chemicals based on their aquatic persistency and mobility. Environ Sci-Proc Imp 19(7):939–955. https://doi.org/10.1039/c7em00158d
71. Deere JR, Streets S, Jankowski MD, Ferrey M, Chenaux-Ibrahim Y, Convertino M, Isaac EJ, Phelps NBD, Primus A, Servadio JL, Singer RS, Travis DA, Seth M, Wolf TM (2021) A chemical prioritization process: applications to contaminants of emerging concern in freshwater ecosystems (Phase I). Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2021.146030
72. Homem V, Silva JA, Ratola N, Santos L, Alves A (2015) Prioritisation approach to score and rank synthetic musk compounds for environmental risk assessment. J Chem Technol Biot 90(9):1619–1630. https://doi.org/10.1002/jctb.4628
73. Sims K (2022) Chemicals of concern: a prioritisation and early warning system for England. ECG Environ Brief 35:19–20. https://doi.org/10.5281/zenodo.7790935
74. Barceló D, Petrovic M (2007) Pharmaceuticals and personal care products (PPCPs) in the environment. Anal Bioanal Chem 387(4):1141–1142. https://doi.org/10.1007/s00216-006-1012-2
75. Baldwin AK, Corsi SR, Stefaniak OM, Loken LC, Villeneuve DL, Ankley GT, Blackwell BR, Lenaker PL, Nott MA, Mills MA (2022) Risk-based prioritization of organic chemicals and locations of ecological concern in sediment from Great Lakes tributaries. Environ Toxicol Chem 41(4):1016–1041. https://doi.org/10.1002/etc.5286
76. Schulze B, Jeon Y, Kaserzon S, Heffernan AL, Dewapriya P, O'Brien J, Ramos MJG, Gorji SG, Mueller JF, Thomas KV, Samanipour S (2020) An assessment of quality assurance/quality control efforts in high resolution mass spectrometry non-target workflows for analysis of environmental samples. TrAC Trends Anal Chem. https://doi.org/10.1016/j.trac.2020.116063
77. Dürig W, Tröger R, Andersson PL, Rybacka A, Fischer S, Wiberg K, Ahrens L (2019) Development of a suspect screening prioritization tool for organic compounds in water and biota. Chemosphere 222:904–912. https://doi.org/10.1016/j.chemosphere.2019.02.021
78. Alsop D, Wilson JY (2019) Waterborne pharmaceutical uptake and toxicity is modified by pH and dissolved organic carbon in zebrafish. Aquat Toxicol 210:11–18. https://doi.org/10.1016/j.aquatox.2019.02.008
79. Li FF, Bao YY, Chen LYJ, Su ZG, Tang YS, Wen DH (2023) Screening of priority antibiotics in Chinese seawater based on the persistence, bioaccumulation, toxicity and resistance. Environ Int. https://doi.org/10.1016/j.envint.2023.108140
80. Ozgun OK, Basak B, Eropak C, Abat S, Kirim G, Girgin E, Hanedar A, Gunes E, Citil E, Görgün E, Gomec CY, Babuna FG, Ovez S, Tanik A, Ozturk I, Kinaci C, Karaaslan Y, Gucver SM, Siltu E, Orhon AK (2017) Prioritization

methodology of dangerous substances for water quality monitoring with scarce data. Clean Technol Envir 19(1):105–122. https://doi.org/10.1007/s10098-016-1194-z
81. Zhong MM, Wang TL, Zhao WX, Huang J, Wang B, Blaney L, Bu QW, Yu G (2022) Emerging organic contaminants in Chinese surface water: identification of priority pollutants. Eng-Prc 11:111–125. https://doi.org/10.1016/j.eng.2020.12.023
82. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE (2024) PubChem 2025 update. Nucleic Acids Res 53(D1):D1516–D1525. https://doi.org/10.1093/nar/gkae1059
83. Götz CW, Stamm C, Fenner K, Singer H, Schärer M, Hollender J (2010) Targeting aquatic microcontaminants for monitoring: exposure categorization and application to the Swiss situation. Environ Sci Pollut Res 17(2):341–354. https://doi.org/10.1007/s11356-009-0167-8
84. Chen X, He T, Yang X, Gan Y, Qing X, Wang J, Huang Y (2023) Analysis, environmental occurrence, fate and potential toxicity of tire wear compounds 6PPD and 6PPD-quinone. J Hazard Mater 452:131245. https://doi.org/10.1016/j.jhazmat.2023.131245
85. Hiki K, Asahina K, Kato K, Yamagishi T, Omagari R, Iwasaki Y, Watanabe H, Yamamoto H (2021) Acute toxicity of a tire rubber-derived chemical, 6PPD quinone, to freshwater fish and crustacean species. Environ Sci Technol Lett 8(9):779–784. https://doi.org/10.1021/acs.estlett.1c00453
86. US EPA (2024) Summary of the toxic substances control act. https://www.epa.gov/laws-regulations/summary-toxic-substances-control-act. Accessed 2 Dec 2024.
87. Arnold KE, Brown AR, Ankley GT, Sumpter JP (2014) Medicating the environment: assessing risks of pharmaceuticals to wildlife and ecosystems. Philos Trans R Soc Lond B Biol Sci. https://doi.org/10.1098/rstb.2013.0569
88. Reid MJ, Wennberg AC, Stoll C, Rostkowski P (2025) Database for pikme – efficient prioritization, identification, and quantification of emerging environmental pollutants. Zenodo. https://doi.org/10.5281/zenodo.15647470
89. European Commission (2022) Proposal for a Directive of the European Parliament and of the Council amending Directive 2000/60/EC establishing a framework for Community action in the field of water policy, Directive 2006/118/EC on the protection of groundwater against pollution and deterioration and Directive 2008/105/EC on environmental quality standards in the field of water policy. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52022PC0540.
90. European Commission (2025) Neonicotinoids. https://food.ec.europa.eu/plants/pesticides/approval-active-substances-safeners-and-synergists/renewal-approval/neonicotinoids_en. Accessed 10 March 2025.

Publisher's Note