

Environmental Science Processes & Impacts

rsc.li/espi



ISSN 2050-7887



Cite this: *Environ. Sci.: Processes Impacts*, 2022, **24**, 1133

Fate-directed risk assessment of chemical mixtures: a case study for cedarwood essential oil†

Roxana Sühring, *^{ab} Philipp Mayer, ^c Pim Leonards^d and Matthew MacLeod ^a

The environmental risk assessment of UVCBs (*i.e.*, substances of unknown or variable composition, complex reaction products, or biological materials) is challenging due to their inherent complexity. A particular problem is that UVCBs can contain constituents with unidentified chemical structures and/or have variable composition of constituents from batch to batch. Moreover, the composition of a UVCB in the environment is not the same as that of the UVCB in a product, meaning that a risk assessment based on environmental exposure to the UVCB in a product does not represent the actual environmental risk. Here we propose an *in silico* fate-directed risk assessment framework for UVCBs using cedarwood oil as a case study. The framework uses Monte Carlo simulations and the mass-balance models SimpleTreat and RAIDAR to provide quantitative information on whether unidentified constituents within the physical–chemical property space of a UVCB can be the decisive factor for the environmental risk of the entire UVCB. Thereby the framework provides a robust decision tool to evaluate if a UVCB risk assessment requires additional tests or if the data on known constituents is representative for the risk of the entire UVCB. In the case of cedarwood oil, it could be shown that a risk assessment based on the known constituents (representing around 70% of the overall UVCB by weight) is representative for the environmental risk of the entire UVCB – reducing the need for additional testing and test animals.

Received 10th March 2022
Accepted 31st May 2022

DOI: 10.1039/d2em00103a
rsc.li/espi

Environmental significance

Environmental risk assessment of UVCBs (substances of unknown or variable composition, complex reaction products, or biological materials) is challenging due to the complexity of the substances and the potential for risks from unknown constituents. An important question is: can a risk assessment of UVCBs be conducted based on the known constituents alone? Here, we present a framework based on Monte Carlo simulations and the mass-balance models SimpleTreat and RAIDAR that helps to answer this question. The framework can improve – and for some UVCBs even enable – a risk evaluation and allows risk assessors to critically evaluate the need for additional tests, which should reduce the need for animal testing.

1 Introduction

Substances of unknown or variable composition, complex reaction products, or biological materials (UVCBs) present a considerable challenge to chemical risk assessors.^{1,2} Testing in support of environmental risk assessment of chemicals usually analyses substances that are made up of discrete chemicals that can be characterized and tested individually.³ UVCBs, in contrast, are intrinsically complex mixtures that

cannot readily be tested as separate constituents, moreover their composition can vary between different batches of the same product, and the chemical structure of some components of UVCBs might be simply unknown.^{1,4}

Different approaches have been proposed by different regulatory agencies to confront the risk assessment challenges posed by UVCBs and evaluate their potential environmental risks. Approaches include an assessment based on the entire UVCB, assessment based on “blocks” (*i.e.*, groups of substances with similar structures), and assessment based on individual constituents above a certain threshold in contribution to the composition of the overall UVCB (for example, 0.1%).^{5,6} Each of these approaches reflects a balance between information requirements to support the assessment and uncertainty and/or lack of fidelity in the description of the UVCB in the assessment compared to the actual substance. As such, all have advantages and shortcomings. For example, both the blocking and threshold approaches could ignore individual constituents that theoretically could be a significant contribution to the

^aDepartment for Environmental Science, Stockholm University, Stockholm, Sweden.
E-mail: roxana.suehring@ryerson.ca

^bDepartment of Chemistry and Biology, Toronto Metropolitan University, Toronto, Canada

^cDepartment of Environmental and Resource Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

^dDepartment of Environment & Health, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1018 HV Amsterdam, Netherlands

† Electronic supplementary information (ESI) available. See <https://doi.org/10.1039/d2em00103a>

environmental risk of the entire UVCB even if they are only present in very low concentrations.⁷

An important consideration in environmental risk assessment of UVCBs is that the composition of the UVCB that is present in a product can change significantly before it enters the environment and within the environment and exposed organisms.⁸ Risk assessment of the neat UVCB without considerations of the environmental fate of the components is, therefore, not necessarily representative of the environmental risk of the UVCB in the environment.^{4,9} Risk, hazard and toxicity assessments that consider changes in composition of UVCBs from the chemical product to the environment and exposed organisms are referred to as fate-directed risk assessments in the scientific literature.⁹

Thus, assessing the potential risk of UVCBs in the environment could benefit from new hazard and risk analysis strategies to complement the existing approaches that (a) account for the environmental fate of different UVCB constituents, and (b) confront the presence of constituents whose structure is not fully known by determining what fraction of a UVCB needs to be fully characterized to assess the risk of the entire UVCB mixture.

Here we illustrate *in silico* fate-directed risk assessment for a UVCB (cedarwood oil) that uses measured, predicted, and literature property and hazard data to drive the mass-balance models SimpleTreat¹⁰ and RAIDAR,^{11,12} with Monte Carlo simulation representing possible unidentified constituents. SimpleTreat uses physical-chemical property data and information on biodegradation to predict the efficiency of removal of chemicals in a wastewater treatment plant.¹⁰ RAIDAR uses chemical property data including environmental half-lives and toxicity to calculate the environmental fate (mass balance for emissions to different media), bioaccumulation and risk (risk assessment factors for different species) of chemicals.^{11,12}

The aim of this study was to develop a fate-directed environmental risk analysis for cedarwood oil that exploits recently published information about biodegradation, bioconcentration and toxicity, and to illustrate a general procedure for fate-based risk assessment of UVCBs with unidentified constituents that

confronts the question of whether unidentified constituents could change the conclusions of a risk assessment of the UVCB that is based only on the known constituents.

2 Methods

2.1 Risk assessment of cedarwood oil based on known constituents

To evaluate the environmental risk of cedarwood oil based only on known constituents, property and hazard data was collected for α -cedrene (CAS: 469-61-4), β -cedrene (CAS: 546-28-1), thujopsene (CAS: 470-40-6), cuparene (CAS: 16982-00-6), and cedrol (CAS: 77-53-2), that together make up >70% of a typical batch of a commercial Virginian cedarwood oil mixture.¹³

The property and hazard data were used as model inputs to (1) SimpleTreat¹⁰ to estimate the removal efficiency, change of composition and emissions into the environment for each constituent during wastewater treatment (Fig. 1), and (2) RAIDAR¹¹ to calculate risk assessment factors for each constituent (Fig. 1).

2.1.1 Characterisation, physical-chemical properties, and PBT property data. Cedarwood oil was selected as a case study substance in the European Chemical Council (Cefic) long-range research initiative (LRI) ECO-42 project,¹⁴ which has generated new data on biodegradation,¹⁵ bioaccumulation potential for fish¹³ and whole substance aquatic toxicity.¹⁶ Characterisation data for the oil was provided by its distributor Givaudan UK Ltd and confirmed by gas-chromatography mass spectrometry (GC-MS).¹³ Physical-chemical property data for the individual cedarwood oil constituents was taken from the literature or estimated using EPISuite (Table 1).¹⁷ Toxicity data for *Daphnia pulex* (EC₅₀) for individual constituents were estimated using a model by Hickey and Passino-Reader.^{18,19} No half-lives were available for birds and mammals. Therefore, the measured half-lives for fish were used as an estimate for the half-lives in birds and mammals.

The following property and hazard data was used for the fate-directed risk assessment of the known constituents (Table 1).

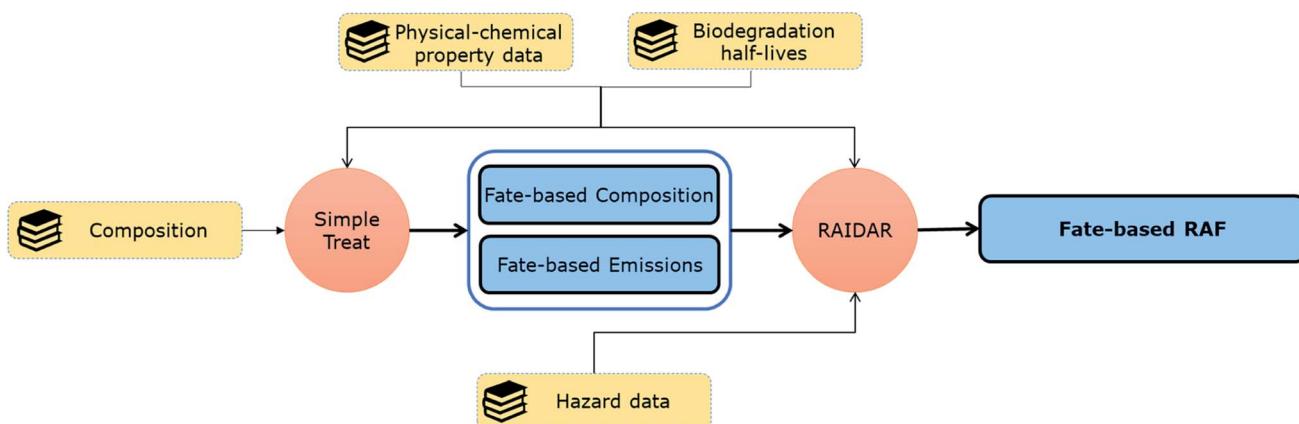


Fig. 1 Schematic of the fate-based risk assessment workflow applied to five known constituents of cedarwood oil that typically make up >70% by weight of the UVCB. Yellow boxes represent the model inputs from literature data and quantitative structure–property relationship (QSPR) models, orange circles represent the models, and blue boxes the model outputs. RAF: risk assessment factor.



Table 1 Weight percent in UVCB product [% weight/weight], molecular mass [g mol⁻¹], half-lives ($T_{1/2}$) [h] for degradation in environmental compartments or biota, water solubility [mg L⁻¹], vapor pressure [Pa], Henry's law constant [Pa m³ mol⁻¹], log octanol–water partitioning coefficient ($\log K_{OW}$), and estimated toxicity (effect concentration 50%) for *Daphnia pulex* (EC₅₀) [μmol L⁻¹] of α -cedrene, β -cedrene, thujopsene, cuparene, and cedrol. All data were estimated using EPISuite¹⁷ unless an alternative source is specified

		α -Cedrene	β -Cedrene	Thujopsene	Cuparene	Cedrol
$T_{1/2}$ [h]	Weight% in UVCB product	26	3.9	19	3.0	22
	Molecular mass [g mol ⁻¹]	204.35	204.35	204.35	202.34	222.37
	Air	2.8	3.5	0.52	27	13
	Water	173 ^a	134 ^a	166 ^a	175 ^a	70 ^a
	Soil	1800	1800	2880	1800	2880
	Sediment	8100	8100	13 000	8100	13 000
	Fish	1272 ^b	1128 ^b	576 ^b	240 ^b	72 ^b
	Birds and mammals	1272	1128	576	240	72
	Water solubility [mg L ⁻¹]	0.15	0.13	0.073	0.22	22
	Vapor pressure [Pa]	3.0	3.0	9.01	0.77	0.070
	Henry's law constant [Pa m ³ mol ⁻¹]	3510 ^e	9639	26 780	712	0.168
	$\log K_{OW}$	5.7	5.8	6.1	6.2	4.3
	EC ₅₀ [μmol L ⁻¹]	0.21 ^c	0.21 ^c	0.27 ^c	0.35 ^c	5.5 ^d

^a Møller *et al.*, 2021.¹⁵ ^b Sühring *et al.*, 2021.¹³ ^c Hickey and Passino-Reader, 1991.¹⁸ ^d REACH registration dossier for Cedrol.²⁰ ^e Copolovici and Niinemets, 2015.³¹

2.1.2 SimpleTreat. SimpleTreat version 4.0 (ref. 10) (available at: <https://www.rivm.nl/en/soil-and-water/simpletreat>) was used to estimate the fate and removal efficiency of the individual cedarwood oil constituents in wastewater treatment plants. The following input data was required for the wastewater treatment simulations:

Molecular mass [g mol⁻¹], octanol–water partitioning coefficient, vapour pressure [Pa], water solubility [mg L⁻¹], and Henry's law constant [Pa m³ mol⁻¹] (Table 1). The chemical class for all constituents was specified as “neutral organic compound”. Based on the biodegradation kinetics measured within the ECO42 project¹⁵ and on previous measurements by Jenner *et al.*,²¹ all constituents were specified to be “inherently biodegradable” in the SimpleTreat model inputs, which implies that they will not be readily removed by biodegradation.

The default SimpleTreat emission scenario (1 kg d⁻¹) was used to model the removal efficiencies and fractions of substance emitted to air and retained in sludge in WWTPs for each individual constituent. These results were used as described below (eqn (1)–(3)) to estimate total emissions of cedarwood oil constituents to the environment to drive the RAIDAR model.

2.1.3 RAIDAR. RAIDAR version 2.02 (ref. 11) (available at: <https://arnotresearch.com/raidar/>) was used to model fate-based risk assessment factors (RAF) for each of the five main cedarwood oil constituents. In RAIDAR, the risk assessment factors (RAF) are derived from integrating information about persistence (P), bioaccumulation (B), toxicity (T), and an estimated emission rate (E_A) for specific substances.^{11,12}

The RAF is the ratio between the emission rate (E_A) of the substance and its critical emission rate (E_C) to induce a “critical” internal concentration in an organism at which toxic effects are expected.^{11,12} The “critical” internal concentration is approximated as the product of the concentration causing acute lethality in 50 percent of a population [mmol L⁻¹] and the bioconcentration factor (BCF, L kg⁻¹).¹²

$$\text{RAF} = \frac{E_A}{E_C} \quad (1)$$

where a RAF ≥ 1 indicates that the actual emissions exceed the critical emission necessary to induce internal concentrations in the organism high enough to produce toxic effects – indicating a risk. Importantly, differences in the RAF between substances or constituents can be used to investigate the relative risk of compounds.

The following input data were used for the RAF estimates for α -cedrene, β -cedrene, thujopsene, cuparene, and cedrol, respectively:

(A) Physical–chemical property data and hazard data presented in Table 1.

(B) Emission estimates.

To estimate emissions, we adopted a conservative scenario in which the maximum of the range of reported production and importation of Virginian cedarwood oil for the European Union (EU) was emitted. According to the REACH registration for cedarwood oil the total import and production of Virginian cedarwood oil is up to 1000 tons per annum.²² The 1000 tons per annum are for the entire EU region which is 45 times bigger area than the modelled region in RAIDAR. Therefore, the total emissions were divided by 45 resulting in a total emission of 22.22 tons per annum for the RAIDAR region.

Based on the typical use of cedarwood oil as fragrance material in personal care products, aroma therapy, and cleaning products²³ we assumed that 50% of the cedarwood oil is emitted into the air and 50% down the drain into the wastewater stream.

For the fate-based risk assessment, the emitted quantity for the individual compounds was calculated as follows:

Emission into air. 50% of the total emitted cedarwood oil (11.11 tons per annum) with a contribution of the known constituents of 26% α -cedrene, 4% β -cedrene, 19% thujopsene, 3% cuparene, and 22% cedrol. Additional emissions into air



during wastewater treatment were also considered based on $f_{air,ST}(x)$, the fraction of constituent x that is predicted by SimpleTreat to be emitted into air during the wastewater treatment process. Thus the total emission to air of each constituent (x) is ($\sum E_{air}(x)$ [tons]):

$$\sum E_{air}(x) = 0.5m_{total}(x) + f_{air,ST}(x) \times 0.5m_{total}(x) \quad (2)$$

where $m_{total}(x)$ is the total mass of the constituent x calculated from the composition of cedarwood oil and the REACH production/import estimates.

For the emissions into water, the 50% total cedarwood oil emissions were multiplied by the effluent emissions estimated using SimpleTreat and the final amount for each constituent was calculated based on its estimated proportion in the mixture following wastewater treatment:

$$\sum E_{water}(x) = f_{water,ST}(x) \times 0.5m_{total}(x) \quad (3)$$

$\sum E_{water}(x)$: total emission of constituent x into water [tons], $f_{water,ST}(x)$: fraction of constituent x that is predicted to be emitted in the effluent of the wastewater treatment process by SimpleTreat.

As a worst-case scenario, we assumed that all sludge from the wastewater treatment process would be applied onto agricultural land as fertilizer leading to 100% emissions of the cedarwood oil in the sludge into the soil. The emissions into soil for the individual cedarwood oil constituents were thus calculated as follows:

$$\sum E_{soil}(x) = f_{soil,ST}(x) \times 0.5m_{total}(x) \quad (4)$$

$\sum E_{soil}(x)$: total emission of constituent x into soil [tons], $f_{soil,ST}(x)$: fraction of constituent x that is predicted to be removed into sewage sludge during the wastewater treatment process.

In addition, a reference scenario representing a lack of wastewater treatment was run in which equal emissions of 11.11 tons per annum of cedarwood oil into air and water were assumed with no direct emissions into the soil.

The emissions into air, water, and soil used as input data in the RAIDAR WWTP scenario and reference scenario for each constituent are presented in Table S1.†

2.2 Assessing the risk of unidentified constituents

To probabilistically estimate the risk of unidentified constituents we developed a Monte Carlo analysis framework (Fig. 2). Firstly, the physical-chemical properties of the known constituents are evaluated along with knowledge of the likely structures of unidentified constituents, and, if possible, this information is used to define a model input parameter space of the unidentified constituents. If a property space of the unidentified constituents cannot be estimated from the chemical property space of the known constituents, then additional characterization and testing of the UVCB is required to enable such estimation. Secondly, a similar evaluation of hazard data for the known constituents is conducted to estimate a toxicity

space for unidentified constituents from the toxicity of known constituents. Again, if a distribution of possible hazard values for the unidentified constituents cannot be defined then additional testing and characterization of the UVCB is required.

If distributions of physical-chemical properties and toxicity hazard of unidentified constituents can be estimated from the known constituents, then fate-based risk assessment for random Monte Carlo realizations of hypothetical combinations of properties representing unidentified constituents is conducted (Fig. 2).

Defining the boundaries of the physical-chemical and hazard property space of unidentified constituents requires expert judgement and should be informed by knowledge of the overall composition of the UVCB and the structural diversity that is expected in the mixture. The constraints of the physical-chemical property space will differ significantly based on the synthesis or production process of individual UVCBs. For example, plant extracts, such as essential oils, can be assumed to contain predominantly terpenes, sesquiterpenes, and chemically related compounds, while complex synthesis products can contain a variety of compounds that are structurally very different – which means that more detailed characterization data will be needed. For hazard, the simplest case is when all components of the UVCB can be assumed to act as baseline toxicants, which is the case for our case study UVCB, cedarwood oil.

In the following subsection the methods for each step in the fate-based risk assessment workflow (Fig. 2) for unidentified constituents are presented in more detail.

2.2.1 Defining the property and hazard space of unidentified constituents. Our assessment approach relies on estimating probability distributions of physical-chemical properties, environmental half-lives, and toxicity of unidentified constituents of the UVCB from the known constituents. These probability distributions are then used to create Monte Carlo realizations of hypothetical constituents with random combinations of properties. Depending on how well the boundaries of the potential properties of the unidentified constituents are constrained, the property space boundaries may extend considerably outside the range of properties of the known constituents to account for structural diversity in the UVCB and data uncertainty.

In case of cedarwood oil, the chemical property space of the UVCB was defined based on the range of physical-chemical property data, environmental half-lives, and ecotoxicity of the five known constituents (Table 1).

Physical-chemical property data differs considerably between the different cedarwood oil constituents – spanning more than 5 orders of magnitude for the most variable property, Henry's law constant (Table 1). Generally, the properties of the polycyclic sesquiterpenes α -cedrene, β -cedrene, and thujopsene were within the same order of magnitude with a $\log K_{ow}$ around 6 (5.7 for α -cedrene to 6.1 for thujopsene), and a water solubility around 0.1 mg L^{-1} (0.073 for thujopsene to 0.15 for α -cedrene) (Table 1). Vapour pressure and Henry's law constant showed slightly larger variability among the polycyclic sesquiterpenes with vapour pressures ranging from 3 Pa for the cedrenes to 9.01 Pa for thujopsene and Henry's law constant ranging from 3510 Pa m^3



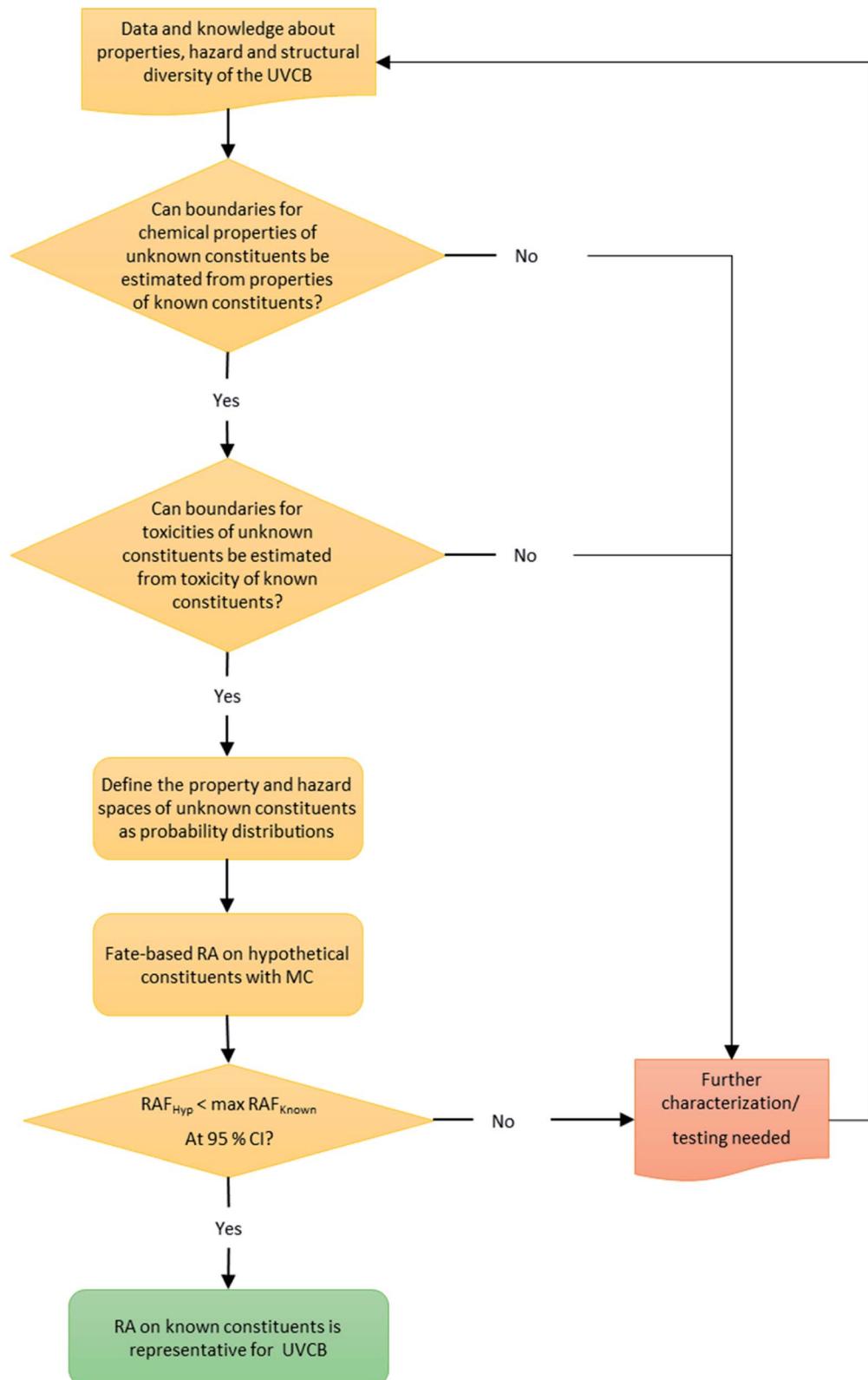


Fig. 2 Workflow for the fate-based risk assessment of unidentified UVCB constituents and decision criteria for whether further characterization/hazard data is needed or if the risk of the known constituents is representative for the entire UVCB based on a Monte Carlo (MC) analysis.

mol^{-1} for α -cedrene to $26\ 780\ \text{Pa m}^3\ \text{mol}^{-1}$ for thujopsene (Table 1). The polycyclic sesquiterpenes were considerably more volatile and less water soluble than the mixed aromatic-cycloaliphatic

cuparene and, especially, the sesquiterpene alcohol cedrol with a water solubility for cuparene and cedrol of $0.22\ \text{mg L}^{-1}$ and



22 mg L⁻¹, and vapor pressure of 0.77 Pa for cuparene and 0.07 Pa for cedrol, respectively (Table 1).

Similarly, the environmental half-lives varied from minutes (thujopsene) to over a day (cedrol) for air and a few days (cedrol) to months (α -cedrene) for half-lives in fish (Table 1). The least variability was observed for half-lives in soil and sediment; however these were EPI-Suite predictions that are very uncertain in themselves.²⁴

Cedarwood oil is prepared by steam distillation, and the five known constituents are expected to represent a large proportion of the variability in structures and properties in the UVCB. To evaluate the physical-chemical property space we conducted an in-depth review of the available regulatory and literature data. The range of observed data informed our general approach to estimating the uncertainty, as well as expert judgement for individual property space boundaries. Generally, for properties with available measured data, a uniform distribution was assumed for each of the properties within the chemical property space with the boundaries of the chemical space for each individual property being the minimum and maximum of the respective property based on the known constituents $\pm 10\%$. For estimated physical-chemical properties an additional 10% (resulting in a total of 20%) were added to account for the uncertainty of the estimates. The resulting value was rounded to the next significant figure. For log K_{ow} the upper boundary was extended to 10 due to the variability of log K_{ow} values for the highly hydrophobic constituent that was observed in the literature. For environmental half-lives a factor 10 was added to account for uncertainty for estimated half-lives. These uncertainties were based on the variability of physical-chemical property data available for the known constituents, as well as the uncertainty used for environmental half-lives used in the OECD P_{ov} and LRT Screening Tool.²⁵ For measured half-lives, the added uncertainty was reduced to the measurement uncertainty. In addition, expert judgement based on the available literature was applied to avoid unreasonably high or low half-lives. The resulting chemical property space for the individual physical-chemical properties, environmental half-lives, and ecotoxicity are presented in the ESI (Table S2†). These probability distributions were randomly sampled to create 2500 hypothetical cedarwood oil constituents with combinations of properties extrapolated from the chemical property space of the known constituents.

2.2.2 Monte Carlo modelling with RAIDAR for unidentified constituents. The fate-based RAIDAR assessment of the unidentified constituents should be conducted on a sufficiently large number of hypothetical constituents to produce replicable results in consecutive Monte Carlo analyses. The emission estimates are informed by the emissions of the total UVCB (assuming 100% emissions as a worst-case scenario), the use patterns to estimate the emissions into air, soil, directly into water, and down the drain into the wastewater stream, as well as the removal and partitioning in a wastewater treatment plant (based on the SimpleTreat analysis). The assumed emission of an unidentified constituent should be \geq the contribution [%] of any unidentified constituent in the UVCB.

For cedarwood oil, the fate-based RAIDAR assessment was conducted for the 2500 hypothetical constituents, which was a sufficient number of iterations to provide replicable results, and using the following emission assumptions:

The contribution of the hypothetical unidentified constituent to the cedarwood oil mixture was assumed to be 1% which we previously estimated was the highest contribution of any unidentified constituent in the characterization of the pure cedarwood oil based on relative peak areas in gas chromatographic analysis of the oil.¹³ As in the fate-based risk assessment of the known constituents, 50% of the used amount was assumed to be emitted down-the-drain into the wastewater stream and 50% into air. Based on the average SimpleTreat results for the known constituents, 50% of the emissions into wastewater was assumed to be eliminated in the primary settler, 5% was assumed to be eliminated *via* the surplus sludge, and 2% were assumed to be biodegraded. The remaining amount was expected to be mostly emitted into the air (97%) while 3% were expected to be emitted in the effluent.

The resulting total emission of a hypothetical unidentified constituent with 1% contribution to the total cedarwood oil UVCB was 0.0015 tons per annum into water, 0.061 tons per annum into soil, and 0.16 tons per annum into air.

The modelled RAFs for different species were used to calculate the 95% confidence level (ranging from the 2.5th percentile to 95.7th percentile) of the potential RAFs of an unidentified constituent within the physical-chemical property space of cedarwood oil.

2.3 Model sensitivity and uncertainty

To test the sensitivity of the model to changes in specific input parameters, the contribution of each input parameter to the variance of the predicted RAFs was calculated based on the 2500 hypothetical cedarwood oil constituents within the chemical property space of the known constituents. As presented above, 10% uncertainty was assumed for the boundaries of the chemical property space based on measurement results. For estimated physical-chemical properties an additional 10% uncertainty were added. For estimated environmental half-lives a factor of 10 was assumed to account for the uncertainty.

3 Results

3.1 Risk of known constituents

3.1.1 Removal in wastewater treatment plants. SimpleTreat modelling indicated that all five known cedarwood oil constituents are removed from wastewater with over 90% efficiency. All known cedarwood oil constituents apart from cedrol had effluent emissions $< 3\%$. Cedrol had predicted 6% emissions *via* the effluent consistent with its lower log K_{ow} compared to the other constituents. Cedarwood oil constituents were estimated to be predominantly eliminated through partitioning into the sludge (Table S3†). Volatilization was the second highest removal process (Table S3†). Cedrol was the only constituent that was estimated to be predominantly removed from the wastewater through volatilization – with adsorption onto sludge



being the second highest removal pathway. This was again consistent with the lower $\log K_{ow}$ compared to the other constituents (Table S3†). Between 23% (cuparene) and 65% (cedrol) were estimated to be emitted from the WWTP into air.

3.1.2 Emissions of cedarwood oil constituents from the wastewater treatment plant. The total estimated emissions of cedarwood oil constituents from wastewater treatment plants into water, soil, and air were >90% for all constituents (Table S3†). Thus, less than 10% of the known constituents were removed by biodegradation, which is consistent with the compounds being “inherently biodegradable” rather than “readily biodegradable” in the SimpleTreat inputs. Of these emissions an average of 58% were through the application of sludge onto soil, 35% were released into the air, and 3% into water (Table S3†). The differences in removal within the wastewater treatment plant for the individual constituents resulted in a relative increase of the contribution of cedrol to the overall cedarwood oil mixture compared to the pure oil in water and air, whereas thujopsene and the cedrenes were reduced compared to their contribution in pure oil (Fig. 3).

3.1.3 The relative amounts of each constituent calculated by SimpleTreat were used as the input for RAIDAR

Estimated environmental fate of cedarwood oil constituents. Modelled emissions into air and water in RAIDAR for most constituents were predicted to partition to >80% into the sediment. Cedrol was the only constituent that was predicted to remain in the water phase (38.4%) or in soil (35.6%). Less than 1% of the constituents were predicted to remain in the air, except for cedrol (<6%) and cuparene (<2%) (Table 2).

For the cedarwood oil emitted onto soil *via* sludge application, close to 100% was predicted to remain in the soil for all constituents (Table S4†).

Estimated environmental risk

Reference scenario: equal emissions into water and air, no wastewater treatment. In the reference scenario with equal emissions into air and water and no wastewater treatment, the modelled RAFs ranged from 3.4×10^{-4} for dairy cows exposed

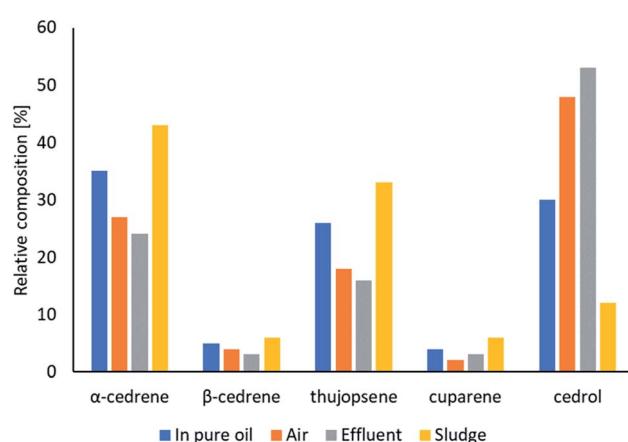


Fig. 3 Changes in the relative composition [%] of cedarwood oil after WWTP in air (orange), effluent (grey) or sludge (yellow) compared to the pure oil (blue).

Table 2 Predicted distribution [%] of the individual cedarwood oil constituents between air, water, sediment, and soil in the fate-based and reference scenario with emissions into air and water

Constituent	% Air	% Water	% Sediment	% Soil
α-Cedrene	0.37	10.9	88.7	0.0159
β-Cedrene	0.442	9.33	90.2	0.0173
Thujopsene	0.0297	4.64	95.3	1.13×10^{-3}
Cuparene	1.26	5.38	93.3	0.0602
Cedrol	5.8	38.4	20.2	35.6

to cedrol to an RAF of 1.9 for aquatic mammals exposed to α -cedrene (Fig. 4, Table S5†). All constituents apart from cedrol were predicted to pose the highest risk for aquatic mammals.

Fate-directed scenario: emissions into water and air following wastewater treatment. In the fate-directed scenario with emissions to water being reduced and partially re-directed to air and soil by wastewater treatment, the estimated RAF remained well below 1 for all cedarwood oil constituents (Fig. 4, Table S6†). For emissions into air and water, α -cedrene and β -cedrene had the highest predicted RAFs with 0.11, followed by thujopsene (0.040), cuparene (0.015), and cedrol (6.6×10^{-4}) (Fig. 4, Table S6†). Similar to the risk estimates without wastewater treatment, the model predicted the highest risk for aquatic mammals for all constituents apart from cedrol which had the highest RAF for cows (Table S6†).

The maximum RAF for emissions into soil was 9.7×10^{-3} , an order of magnitude below the predicted RAFs from emissions into air and water. The most vulnerable species were modelled to be Avian omnivores (for α -cedrene and cuparene) and terrestrial invertebrates (for β -cedrene, thujopsene, and cedrol) (Fig. 4, Table S7†).

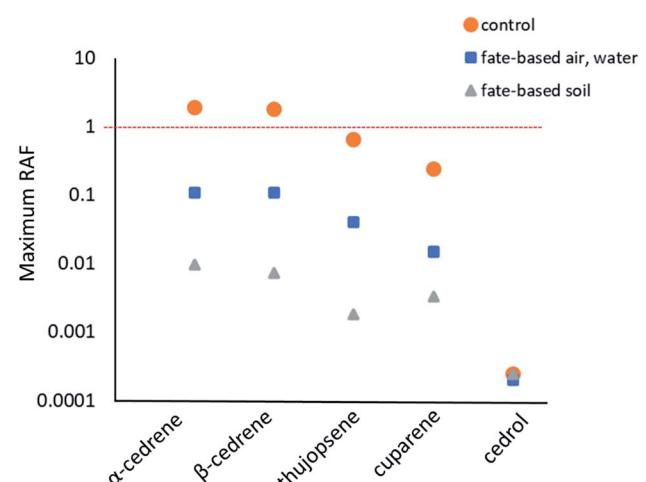


Fig. 4 Maximum risk assessment factors (RAF) for α -cedrene, β -cedrene, thujopsene, cuparene, and cedrol from emission into air, water (orange), and soil (grey) with wastewater treatment and into air and water without wastewater treatment (blue). The red dotted line indicates the cut-off for expected risk.



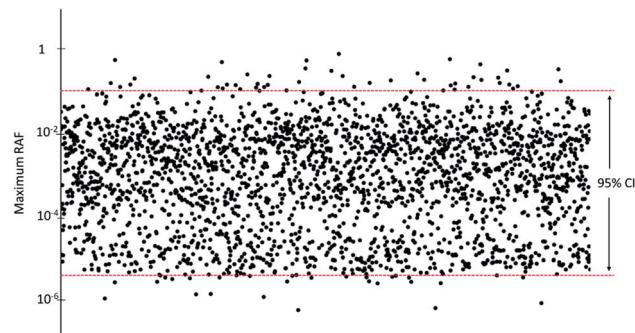


Fig. 5 Maximum risk assessment factors (RAF) for 2500 hypothetical unidentified constituents of cedarwood oil with a 1% contribution to the total weight of the cedarwood oil mixture. The dotted red lines present the 2.5th percentile to 97.5th percentile.

3.2 Environmental risk of unidentified constituents

The predicted RAF for a hypothetical constituent with 1% contribution to the cedarwood oil mixture and properties randomly sampled from the chemical property space estimated from the known constituents was between 3.9×10^{-11} for root vegetables and 0.097 for aquatic mammals at a 95% confidence level. The maximum RAF was 0.80 for aquatic mammals (Fig. 5, Table S8†). In only 49 of the 2500 Monte Carlo iterations (2%) did the highest RAF of the hypothetical constituent exceed that of the known constituents (Fig. 5), and in no case did the RAF exceed 1.

4 Discussion

4.1 Environmental risk from cedarwood oil

The results indicate a potential risk from α -cedrene and β -cedrene for aquatic mammals in the very conservative reference scenario that assumed that all produced and imported cedarwood oil in the EU would be emitted into the environment without wastewater treatment. In the more realistic, but still conservative fate-based scenario that assumed that all produced and imported cedarwood oil in the EU would be emitted but included wastewater treatment for emissions into water the RAF was reduced by a factor of 10 for most constituents (Fig. 4). This indicated that the environmental risk of the cedarwood oil constituents was driven by emissions into water rather than emissions into the air or by applying sewage sludge onto soil, because those emissions were not reduced in the fate-based scenario.

Cedarwood oil is used as fragrance material in personal care products, pet care products, aroma therapy, cleaning products, as well as insect repellent and insecticide (NIH, 2002). Average amounts in the respective products range from <1% in most personal care products to 5% in laundry and fabric treatment (Table S9†).^{23,26} Moreover, pure cedarwood essential oil can be used in aroma therapy.²⁷ The use of cedarwood oil in leave on or rinse-off products supports the assumption that cedarwood oil emissions into water will occur overwhelmingly down-the-drain which means that it will enter a wastewater treatment plant,

which will reduce the potential environmental risk considerably. However, pulp mills have been suspected to be sources of direct cedarwood emissions into water.¹⁹

That the water emissions drive the overall risk of cedarwood oil constituents explains why aquatic mammals were modelled to have the highest RAFs. Aquatic mammals generally have higher lipid contents or occupy higher trophic levels than most fish species, resulting in a higher accumulation of hydrophobic chemicals such as the cedarwood oil constituents.²⁸ Cedrol has a lower log K_{ow} than the other known constituents investigated in this study, resulting in a lower modelled accumulation in aquatic mammals and consequently a lower RAF.

4.2 Is the risk assessment based on the known constituents representative for the entire UVCB?

Our proposed workflow for the fate-based risk assessment of the unidentified constituents has three important decision criteria that have to be evaluated to decide whether the risk assessment based on the known constituents is representative for the entire UVCB (Fig. 2). The decision criteria outlined below ensure that the framework can be applied for UVCBs other than cedarwood oil or other essential oils.

(a) Can a distribution of physical-chemical properties that represents the unidentified n constituents be estimated from the chemical properties of the known constituents and other information about the UVCB? This decision criterion was established to ensure that the hypothetical unidentified constituents created for the Monte Carlo analysis are representative for the range of substances that could be present in the UVCB. Based on the available characterization data for cedarwood oil,¹³ it was assumed that all unidentified constituents would be terpenes or sesquiterpenes, meaning that none of the unidentified constituents was expected to fall considerably outside the chemical property space of the known constituents. However, this criterion needs to be carefully evaluated for UVCBs with more diverse physical-chemical property spaces.

(b) Can a distribution of toxicity that represents the unidentified constituents be estimated from the toxicity of the known constituents? This decision criterion was established to ensure that hazards such as excess toxicity are ruled out before proceeding to the fate-based risk assessment of the unidentified constituents. Baseline toxicity and thus absence of excess toxicity was supported for cedarwood oil based on the whole UVCB toxicity tests that were conducted within the ECO-42 project.¹⁶ For a general application of our framework for UVCB risk assessment, potential excess toxicity needs to be evaluated carefully and additional experiments might be necessary particularly with regards to potential endocrine disruptive effects or carcinogenicity.²⁹

The evaluation of decision criteria 1 and 2 indicated that no further characterization data was needed to perform the fate-based risk assessment of the unidentified constituents.

(c) Does the predicted RAF for the unidentified constituents exceed the maximum RAF of the known constituents at a 95% confidence level? The main question the fate-based risk assessment of the unidentified constituents tries to answer is

whether an unidentified constituent could present a higher risk than any of the known constituents. In case of cedarwood oil, we knew from the characterization data that none of the unidentified constituents would have a contribution of $>1\%$ to the oil mixture.¹³ Therefore we assessed the risk of a hypothetical unidentified constituent with 1% contribution. Based on the results, 2% of the 2500 hypothetical constituents had an RAF that exceeded the RAF of α -cedrene and β -cedrene (the known constituents with the highest predicted RAFs) (Fig. 4, 5, Tables S6 and S8†). At a 95% confidence level the RAF of all unidentified constituents was below the RAF of α -cedrene and β -cedrene (Tables S6 and S8†).

4.3 Risk drivers

The octanol–water partitioning coefficient was identified as the main determinant for the predicted RAFs with a contribution to variance of $>30\%$ (Fig. S1†). The second highest contributor to

the RAF variance was the toxicity with around 4% (Fig. S1†). However, it was interesting to note that hypothetical constituents with predicted RAFs > 0.01 had consistently higher $\log K_{\text{OW}}$, toxicity, persistence, and bioaccumulation potential than hypothetical constituents with lower predicted RAFs, while also having higher predicted water solubility and lower vapor pressure (Fig. 6). While the combination of high K_{OW} , toxicity, persistence, and bioaccumulation potential are not surprising, the high apparent water solubility calls into question whether the combination of properties predicted to lead to RAFs > 0.01 are realistic.

Solubility and K_{OW} are highly correlated. Therefore, the “subcooled liquid solubility (S_L)” can serve as an upper boundary for possible exposure.³⁰ A regression of $\log K_{\text{OW}}$ against S_L can then be used to determine combinations of physical–chemical property combinations that are not feasible.

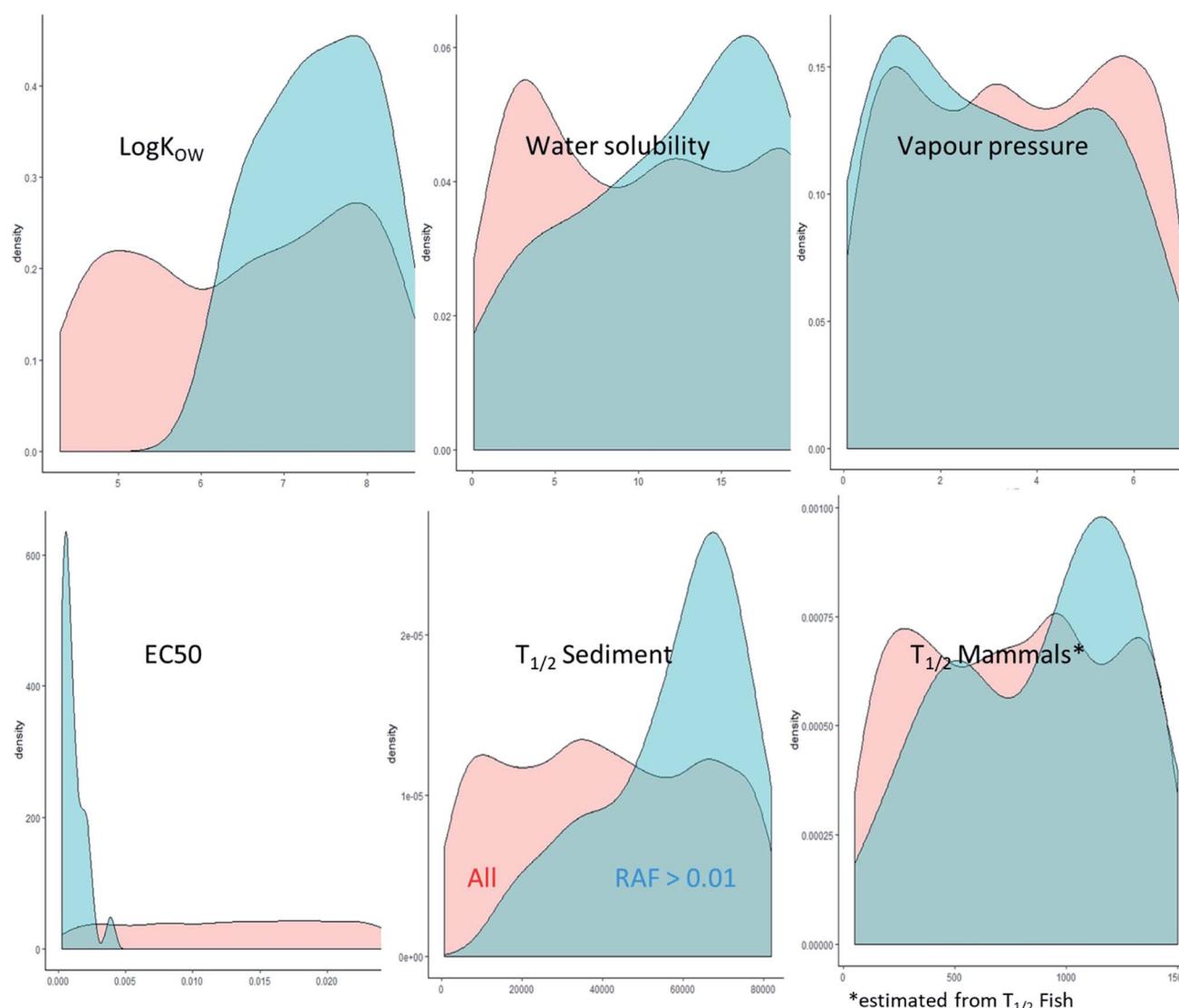


Fig. 6 Prevalence of $\log K_{\text{OW}}$, water solubility, vapour pressure, EC_{50} , $T_{1/2}$ sediment, and $T_{1/2}$ mammals in all hypothetical constituents (red) compared to hypothetical constituents with an estimated RAF > 0.01 (blue).



To evaluate the highest RAF based on a realistic combination of physical-chemical properties, we removed all hypothetical constituents with an unrealistically high water solubility for the respective K_{ow} . This led to an exclusion of 415 of the 2500 hypothetical constituents. The exclusion of predicted RAFs with unrealistically high water solubility for the respective K_{ow} reduced the 95 percentile RAF and median predicted RAF for aquatic mammals and avian scavengers by an order of magnitude (Table S10†) which meant that the difference in RAF between unidentified constituents and known constituents exceeded the input data uncertainty. This provided strong evidence that the risk assessment based on the known constituents was representative for the entire cedarwood oil UVCB.

The developed risk assessment strategy is a robust decision tool to evaluate whether a UVCB risk assessment requires additional tests or if the data on known constituents is representative for the risk of the entire UVCB. As such the framework can reduce unnecessary testing – reducing costs and the need for test animals – as well as highlight data gaps in UVCB risk assessments. The developed risk assessment is based on freely available models, making it widely accessible.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This research was funded by the European Chemical Industry Council Long-Range Research Initiative (CEFIC-LRI) project ECO 42: Fate-directed toxicity testing and risk assessment of UVCBs.

References

- 1 Z. Wang, H. Wiesinger and K. Groh, Time to Reveal Chemical Identities of Polymers and UVCBs, *Environ. Sci. Technol.*, 2021, **55**, 14473–14476.
- 2 A. Lai, A. M. Clark, B. I. Escher, M. Fernandez, L. R. McEwen, Z. Tian, Z. Wang and E. L. Schymanski, The Next Frontier of Environmental Unknowns: Substances of Unknown or Variable Composition, Complex Reaction Products, or Biological Materials (UVCBs), *Environ. Sci. Technol.*, 2022, DOI: [10.1021/ACS.EST.2C00321](https://doi.org/10.1021/ACS.EST.2C00321).
- 3 ECHA, *Guidance on Information Requirements and Chemical Safety Assessment*, ECHA, <https://echa.europa.eu/guidance-documents/guidance-on-information-requirements-and-chemical-safety-assessment>, accessed 13 January 2022.
- 4 S. D. Dimitrov, D. G. Georgieva, T. S. Pavlov, Y. H. Karakolev, P. G. Karamertzanis, M. Rasenberg and O. G. Mekenyanyan, UVCB substances: Methodology for structural description and application to fate and hazard assessment, *Environ. Toxicol. Chem.*, 2015, **34**, 2450–2462.
- 5 Concawe, PETROTOX, Concawe, <https://www.concawe.eu/reach/petrotex/>, accessed 13 January 2022.
- 6 S. S. Kutsarova, D. G. Yordanova, Y. H. Karakolev, S. Stoeva, M. Comber, C. B. Hughes, E. Vaiopoulou, S. D. Dimitrov and O. G. Mekenyanyan, UVCB substances II: Development of an endpoint-nonspecific procedure for selection of computationally generated representative constituents, *Environ. Toxicol. Chem.*, 2019, **38**, 682–694.
- 7 G. Ellis, Special Issue 40th ISEO: toxicological challenges for essential oils in REACH, *Flavour Fragrance J.*, 2010, **25**, 138–144.
- 8 M. MacLeod, T. E. McKone, K. L. Foster, R. L. Maddalena, T. F. Parkerton and D. Mackay, Applications of Contaminant Fate and Bioaccumulation Models in Assessing Ecological Risks of Chemicals: A Case Study for Gasoline Hydrocarbons, *Environ. Sci. Technol.*, 2004, **38**, 6225–6233.
- 9 D. Salvito, M. Fernandez, K. Jenner, D. Y. Lyon, J. de Knecht, P. Mayer, M. MacLeod, K. Eisenreich, P. Leonards, R. Cesnaitis, M. León-Paumen, M. Embry and S. E. Déglis, Improving the Environmental Risk Assessment of Substances of Unknown or Variable Composition, Complex Reaction Products, or Biological Materials, *Environ. Toxicol. Chem.*, 2020, **39**, 2097–2108.
- 10 J. Struijs, *SimpleTreat 4.0: A Model to Predict Fate and Emission of Chemicals in Wastewater Treatment Plants Background Report Describing the Equations*, RIVM Rapp, 2014.
- 11 J. A. Arnot, D. Mackay, E. Webster and J. M. Southwood, Screening level risk assessment model for chemical fate and effects in the environment, *Environ. Sci. Technol.*, 2006, **40**, 2316–2323.
- 12 J. A. Arnot and D. Mackay, Policies for chemical hazard and risk priority setting: Can persistence, bioaccumulation, toxicity, and quantity information be combined?, *Environ. Sci. Technol.*, 2008, **42**, 4648–4654.
- 13 R. Sühring, C. E. Chen, M. S. McLachlan and M. Macleod, Bioconcentration of cedarwood oil constituents in rainbow trout, *Environ. Sci.: Processes Impacts*, 2021, **23**, 689–698.
- 14 CEFIC, ECO42: *Fate-directed Toxicity Testing and Risk Assessment of UVCBs*, Cefic-Lri, <http://cefic-lri.org/projects/eco42-fate-directed-toxicity-testing-and-risk-assessment-of-uvcbs/>, accessed 13 January 2022.
- 15 M. T. Møller, H. Birch, K. K. Sjøholm, R. Hammershøj, K. Jenner and P. Mayer, Biodegradation of an essential oil UVCB – Whole substance testing and constituent specific analytics yield biodegradation kinetics of mixture constituents, *Chemosphere*, 2021, **278**, 130409.
- 16 L. N. Trac, K. K. Sjøholm, H. Birch and P. Mayer, Passive Dosing of Petroleum and Essential Oil UVCBs – Whole Mixture Toxicity Testing at Controlled Exposure, *Environ. Sci. Technol.*, 2021, **55**, 6150–6159.
- 17 US EPA, *EPI Suite™-Estimation Program Interface*, US EPA, <https://www.epa.gov/tscascreening-tools/epi-suitetm-estimation-program-interface>, accessed 13 January 2022.
- 18 J. P. Hickey and D. R. Passino-Reader, Linear solvation energy relationships: ‘rule of thumb’ for estimation of variable values, *Environ. Sci. Technol.*, 2002, **25**, 1753–1760.



19 D. R. Passino-Reader, J. P. Hickey and L. M. Ogilvie, Toxicity to *Daphnia pulex* and QSAR predictions for polycyclic hydrocarbons representative of Great Lakes contaminants, *Bull. Environ. Contam. Toxicol.*, 1997, **59**, 834–840.

20 REACH, *Registration Dossier for Cedrol*, <https://echa.europa.eu/substance-information/-/substanceinfo/100.000.942>, accessed 13 January 2022.

21 K. J. Jenner, G. Kreutzer and P. Racine, Persistency assessment and aerobic biodegradation of selected cyclic sesquiterpenes present in essential oils, *Environ. Toxicol. Chem.*, 2011, **30**, 1096–1108.

22 REACH, *Registration Dossier for Cedarwood Oil (Virginian)*, <https://echa.europa.eu/substance-information/-/substanceinfo/100.077.582>, accessed 13 January 2022.

23 B. Safford, A. M. Api, C. Barratt, D. Comiskey, G. Ellis, C. McNamara, C. O'Mahony, S. Robison, J. Rose, B. Smith and S. Tozer, Application of the expanded Creme RIFM consumer exposure model to fragrance ingredients in cosmetic, personal care and air care products, *Regul. Toxicol. Pharmacol.*, 2017, **86**, 148–156.

24 M. Matthies and S. Beulke, Considerations of temperature in the context of the persistence classification in the EU, *Environ. Sci. Eur.*, 2017, **29**, 1–11.

25 F. Wegmann, L. Cavin, M. MacLeod, M. Scheringer and K. Hungerbühler, The OECD software tool for screening chemicals for persistence and long-range transport potential, *Environ. Model. Softw.*, 2009, **24**, 228–237.

26 A. Williams, *The Chemical and Products Database (CPDat) MySQL Data File*, https://epa.figshare.com/articles/dataset/The_Chemical_and_Products_Database_CPDat_MySQL_Data_File/5352997, accessed 13 May 2022.

27 Healthline, *Cedarwood Essential Oil Benefits, Uses, and Precautions*, <https://www.healthline.com/health/cedarwood-essential-oil#benefits>, accessed 13 January 2022.

28 B. M. Braune, P. M. Outridge, A. T. Fisk, D. C. G. Muir, P. A. Helm, K. Hobbs, P. F. Hoekstra, Z. A. Kuzyk, M. Kwan, R. J. Letcher, W. L. Lockhart, R. J. Norstrom, G. A. Stern and I. Stirling, Persistent organic pollutants and mercury in marine biota of the Canadian Arctic: an overview of spatial and temporal trends, *Sci. Total Environ.*, 2005, **351**–352, 4–56.

29 C. A. Ferraz, M. R. Pastorinho, A. Palmeira-de-Oliveira and A. C. A. Sousa, Ecotoxicity of plant extracts and essential oils: A review, *Environ. Pollut.*, 2022, **292**, 118319.

30 P. Mayer and S. N. Schmidt, Comment on “Assessing Aromatic-Hydrocarbon Toxicity to Fish Early Life Stages Using Passive-Dosing Methods and Target-Lipid and Chemical-Activity Models”, *Environ. Sci. Technol.*, 2017, **51**, 3584–3585.

31 L. Copolovici and U. Niinemets, Temperature dependencies of Henry's law constants for different plant sesquiterpenes, *Chemosphere*, 2015, **138**, 751–757.

