Evaluation, control and Mitigation of the EnviRonmental impacts of shippinG Emissions (EMERGE)

Deliverable 2.3,

"Report on scrubber water whole effluent toxicity testing, at different geographical regions"

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 874990. The content of the deliverable reflects only the author's view and CINEA is not responsible for any use that may be made of the information it contains.

Grant Agreement No.	874990
Project acronym	EMERGE
Project full title	Evaluation, control and Mitigation of the EnviRonmental impacts of shippinG Emissions
Dissemination level	Public
Due date of deliverable	M30
Actual submission date	M33
Deliverable name	D2.3. "Report on scrubber water whole effluent toxicity testing, at different geographical regions"
Type	Report
Status	Final
WP contributing to the deliverable	WP2
Author(s)	Kerstin Magnusson (IVL) & Maria Granberg (IVL)
Other Contributor(s)	Marco Picone, (UV), Annamaria Volpi Ghirardini (UV), Lina Zapata Restrepo (UoS), Malcolm Hudson (UoS), Ian Williams (UoS), Chiau Yu Chen (IVL), Ana Ré (UAV), Nelson Abrantes (UAV), Savvas Genitsaris (AUTH), Polyxeni Kourkoutmani (AUTH), Natassa Stefanidou (AUTH), Maria Moustaka-Gouni (AUTH), Anna Lunde Hermansson (Chalmers), Erik Ytreberg (Chalmers), Ida-Maja Hassellöv (Chalmers)
Reviewer(s)	Elisa Giubilato (UV), Manuel Aghito (MET)
Keywords	Scrubber effluent, ecotoxicology, effects, LOEC, NOEC, EC10, risk assessment, polycyclic aromatic hydrocarbons (PAHs), alkylated PAHs, metals

	Name	Date
Verification Final Draft by WP leader	**Mira Petrović**	**31.10.2022**
Check before upload by coordinator	**Jukka-Pekka Jalkanen**	**31.10.2022**

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AIS	Automated Identification System
CEGQs	Canadian Environmental Quality Guidelines
CRED	Criteria for Reporting and Evaluating Ecotoxicity Data
DNV GL	Det Norske Veritas and Germanischer Lloyd
EC10; EC50	Effective Concentration 10%; Effective Concentration 50%
EQS	Environmental Quality Standards
HFO	Heavy Fuel Oil
IMO	International Maritime Organization
HGO	Heavy Gas Oil
LC10; LC50	Lethal Concentration 10%; Lethal Concentration 50%
LOD/LOQ	Limit of Detection/ Limit of Quantification
LOEC	Lowest Observed Effect Concentration
MARPOL	International Convention for the Prevention of Pollution from Ships
nMDS	Non-metric multi-dimensional scaling
NOEC	No Observed Effect Concentration
PAC	Polycyclic Aromatic Compound
PAH	Polycyclic Aromatic Hydrocarbon
PSSA	Particularly Sensitive Sea Areas
SECA	Sulphur Emission Control Areas
US-EPA	United States Environmental Protection Agency

EXECUTIVE SUMMARY

Scrubber water, that is, wash water from "scrubbing" of exhaust gases from ships, is a new gateway for pollution to the sea, introduced as a response to the stricter sulphur regulation rules approved and adopted by the International Maritime Organization (IMO) in 2020. The decision to accept ship-based scrubber systems instead of the ships switching to cleaner fuel was made without any serious risk assessment of the effects this would have on the marine ecosystems. As a response to the "sulphur cap" the number of ship scrubbers has increased from 242 in 2015 to 4737 in 2022 (DNV GL retrieved September 23, 2022), and very large volumes of combustion pollutants, mainly combustion particles, oil related compounds and metals, are released directly into the sea.

In the EMERGE project, five European universities and research institutes have carried out experiments where the impact of scrubber water has been tested on a range of marine phytoplankton and planktonic invertebrates. The invertebrates included both species living their entire life in the water column, the copepods *Acartia tonsa* and *Calanus helgolandicus*, and planktonic eggs and larvae from animals spending the adult part of their life cycle on the bottom of the sea, the sea urchins *Strongylocentrotus droebachiensis* and *Paracentrotus lividus*, the mussels *Mytilus edulis* and *Mytilus galloprovincialis* and the polychaete *Sabellaria alveolate*.

Toxic effects of scrubber water were detected at considerably lower concentrations than previously reported. The most sensitive of the analysed endpoints was fertilisation of sea urchin eggs, where a statistically significant disturbance was observed already at concentrations of 0.0001% scrubber water, that is 1 mL scrubber water per m^3 of seawater. Scrubber water at concentrations of 0.001% was found to cause significant malformations of larvae of species from three large groups (phyla) of animals; sea urchins, which belong to the phylum Echinodermata, a polychaete belonging to the phylum Annelida and mussels belonging to the phylum Mollusca. For all these the Lowest Observed Effect Concentration (LOECs) could be determined but not the No Observed Effect Concentration (NOEC), since effects were observed already at the lowest concentration tested.

The results being similar in species from different phyla indicates that the adverse effect of scrubber water on invertebrate larvae is of a general character. Effects on copepod egg production was observed at 0.01% scrubber water and larval development at 0,01-0.1% scrubber water (depending on the origin of the scrubber water tested). Two endpoints commonly included in standardised protocols for ecotoxicological tests are mortality of adult individuals of some specified invertebrate species and growth inhibition of microalgae. Both endpoints were also tested within EMERGE. Increased mortality in response to scrubber water was measured in two species of copepods,

juvenile individuals of *Calanus helgolandicus* and adults and larvae from *Acartia tonsa*. For all *Acartia* life stages, mortality was a relatively insensitive endpoint, whereas juvenile *Calanus* were affected at fairly low concentrations. The observed difference might, at least partly, be explained by the fact that the *Acartia* experiments were carried out on animals reared in the laboratory, whereas the *Calanus* were field collected. Growth inhibition of microalgae was tested on both single species of algae and on natural algae communities, consisting of a multitude of species, and in both experimental set-ups the parameter was found to be relatively non-sensitive to the scrubber water in comparison with the other species and life stages tested within EMERGE.

The large difference in sensitivity to scrubber water between the different endpoints that were tested in EMERGE highlights the importance of executing experiments on a variety of species and life stages to identify the sensitive parts of the marine ecosystems. This requires a research approach based on profound biological and ecotoxicological knowledge rather than limiting the scope to only a set of routine standard protocols, using a limited set of species and endpoints. It is thus unfortunate that IMO recommends using standardised protocols only for ecotoxicological testing of scrubber water (IMO 2019). Considering the already existing quality assurance procedures for ecotoxicological data and risk assessment e.g., CRED analysis suggested by EC (2018) and the rigorous scientific peer-review process of scientific journals, the procedure recommended by the IMO may actually produce data of lesser quality and relevance, i.e., data unfit to protect marine life and ecosystems.

ACKNOWLEDGEMENTS

We would like to thank dr. Kent Salo at Chalmers University of Technology for providing the project with scrubber water from the pilot scale scrubber equipment.

1 INTRODUCTION

Heavy fuel oil (HFO) has traditionally been, and still is, the most used of all ship fuels. Combustion of HFO generates emissions with harmful effects on local and regional air quality due to its content of combustion related particles and hazardous chemicals. The emitted particles consist mainly of soot, ash, unburned organic carbon, sulphur particles and elemental carbon (char), and many have metals and polycyclic aromatic compounds (PACs) associated with them (Moldanova et al. 2009). Compounds with the greatest toxic potential are, apart from the metals, PACs and other oil related substances originating from uncombusted HFO or formed during the combustion process.

Air emissions from HFO combustion have caused great concern, and the focus has primarily been on acidification from release of sulphur oxides, eutrophication from release of nitrogen oxides, and hazardous effects on human health from combustion particles (Seddiek and Elgohary 2014, Clear Seas Centre for Responsible Marine Shipping 2022). To improve air quality a limit of the sulphur content in ship fuel, also known as the "sulphur cap", was approved by the International Maritime Organisation (IMO) and entered into force 1 January 2020. The regulation applies to shipping on a global level, and states that the sulphur content in ship fuel must not exceed 0.5% (MARPOL Annex VI, MEPC.280(70)). Certain sea areas, like that of the EU, are considered as extra sensitive to air pollutants and therefore designated as sulphur emission control areas (SECAs), where maximum allowed sulphur content in the fuel is 0.1%. Fuel with higher sulphur content than the IMO restrictions may however be used on ships where an exhaust gas cleaning system, also known as a *scrubber*, is installed. In a scrubber the exhaust gases are continuously sprayed with water, which captures much of the hazardous components. In practice this means that scrubber water in areas classified as SECAs, where restrictions on air emissions are stricter, will have higher concentrations of combustion pollutants in the discharged scrubber water. The scrubbers are designed to operate in a closed or an open mode, where the closed loop system captures and retains more (but far from all) of the exhaust gas content onboard the ship and unload it in port where the "sludge" is later incorporated in cement or asphalt, whereas from the open loop scrubber system the untreated wash water is discharged directly into the sea. There are also hybrid scrubbers that can switch between closed and open mode. The number of ships with scrubbers in operation, or on order, amounts to 4737 in 2022 and will rise to 4897 in 2024 (DNV-GL 2022). Globally, most ships equipped with scrubbers have an open loop model, but in some SECAs the use of closed and hybrid scrubbers is more common.

Some of the SECAs, i.e., areas where restrictions on air emissions are particularly strict, also have a marine environment that is considered particularly sensitive to anthropogenic stressors and therefore designated by IMO as Particularly Sensitive Sea Areas (PSSAs). The definition of a PSSA is "*an area that needs special protection through action by IMO because of its significance for recognized ecological or socio-economic or scientific reasons and which may be vulnerable to damage by international maritime activities*". The irony is hence that the marine ecosystems in sea areas classified as both PSSA and SECA, like for example the Baltic Sea, the North Sea and the Wadden Sea, will receive higher pollutant concentrations via scrubber water than sea areas with no SECA classification. This is simply because shipping emissions are not managed with a holistic approach, and then regulations on air emissions can be implemented without considering subsequent impacts on other environmental compartments, in this case the marine ecosystem.

2 BACKGROUND

The decision to allow the use of scrubbers instead of switching to a "cleaner" fuel as a measure to comply with the new sulphur cap, was based almost exclusively on a concern about terrestrial acidification and human health. However, no risk assessments were carried out on the impact of the discharged wash water from the scrubbers on the marine ecosystems prior to the decision, which must be regarded as a serious neglect. The sea is already under a great pressure from numerous stressors, not least pollutants, deriving from a multitude of sources including deposition directly into the sea, runoff from land, and air deposition on the sea surface. To switch the input of hazardous components in HFO combustion gases to the sea from a more diffuse deposition from air to a direct input to the sea through scrubber water is a dramatic change for marine organisms. The air does not sustain an ecosystem in the same way that terrestrial or aquatic parts of the biosphere do. The main direct effects of emitting HFO combustion gases to air, besides the combustion particles contributing to climate change, will therefore not occur until the hazardous particles and chemicals are deposited and get in touch with organisms in the terrestrial or aquatic ecosystems. At that time the particles and pollutants have dispersed considerably already in the atmosphere, which reduces the risk for acute toxic effects in the water. Discharges of the content in combustion gases via scrubber water to the sea will, on the other hand, come into direct contact with the aquatic ecosystems the moment the water leaves the discharge tube. The water column close to the surface where the scrubber water is discharged, holds a rich and diverse community of planktonic algae and crustaceans, larvae of bottom living organisms like mussels, oysters and many crayfish species, fish larvae and adult fish, and many other species. In areas where ship lanes run close to the coast, also the shallow water ecosystems are at risk. Population effects of the species that are the primary targets of scrubber water pollutants may lead to destabilised marine ecosystems and have consequences on species important for human food supply, such as fish and crayfish (Heath and Lough 2007, Beaugrand and Kirby 2010b, Stige et al. 2011).

The marine ecosystems in the vicinity of ship lanes where scrubber water is discharged hence run a serious risk to encounter toxic concentrations of the content of combusted HFO. It could be argued that the ship lanes cover only limited areas of the sea. However, data from automated identification system (AIS) on ship movement reveal that traffic intensity along many ship lanes is very high, and since the discharge of scrubber water is continuous as long as the ships are operating, the amount of pollutants deposited in these sensitive areas may be substantial (http://www.shiptraffic.net). There is also a risk that busy ship lanes with high concentrations of hazardous components from

HFO combustion act as three-dimensional toxic barriers and disturb the natural and important movements and migrations of pelagic marine organisms, like larvae and fish, between sea areas (Folt and Burns 1999, Pineda et al. 2007, Chan et al. 2018).

The amount of hazardous particles and compounds deposited in the sea via scrubber water depends on the chemical content and volumes of the discharged water. Collection of data and calculations of emission factors from open- and closed loop scrubbers have been done within the EMERGE project (D2.1) but there is still limited data on the quantities of contaminants from scrubber water reaching different sea areas (Ytreberg et al. 2021). The input of polycyclic aromatic hydrocarbons (PAHs, an important group of oil derived pollutants), from shipping to the Baltic Sea has been estimated to derive almost entirely, ≥98%, from scrubber water (Ytreberg et al. 2022), however, data on the contribution of PAHs from shipping relative to other sources, e.g., riverine input and point sources from industrial and urban activities, in the Baltic Sea is still uncertain and suggested to be between 0.4 and 8.9%. It is also important to point out that up to date, chemical analysis on oil related compounds in scrubber water has mainly been restricted to include the United State Environmental Protection Agency (US EPA) 16 PAHs, whereas the true content of these compounds in water from scrubbing HFO combustion fuels most likely includes hundreds of aliphatic compounds, monoaromatics and PACs such as non-substituted-, alkylated- heterocyclic- and other PAHs (e.g., Zhao et al. 2020 and recent data from EMERGE D2.2). The documentation of the complex chemical composition of scrubber water is in sharp contrast to the present legislation on threshold levels for content in scrubber water discharged from ships, which includes but a few PAHs, pH, and turbidity. Threshold levels set by the legislation are hence far from comprehensive enough to protect the marine ecosystems (IMO Resolution MEPC.259(68) 2015 Guidelines for Exhaust Gas Cleaning Systems) (IMO 2015).

For some contaminants present in scrubber water there are available threshold levels that must not be exceeded in marine environments, e.g., the environmental quality standards (EQS) of the Water Framework Directive (2008/105/EC) for coastal water bodies, or thresholds (often the same EQS) set within Descriptor 8 of the Marine Strategy Framework Directive (DIR 2008/56/EC) for coastal, territorial water and open sea. Assessments of the effect of scrubber water once it is released to the sea are still rare, but when they have been done, they have often been based on concentrations of individual compounds and comparisons to these threshold levels. These comparisons are helpful when toxicity data based on the whole water is limited or missing. However, it is important to keep in mind that this approach will only consider the substances that were part of the chemical analyses, while the majority of potentially toxic compounds that are present but not analysed, will be

neglected. Basing the risk assessment on threshold levels for individual compounds will also miss the effect the contaminants have when acting together in a mixture, as would be the case when a scrubber water is discharged into the sea. It is a well-documented fact that the effect on living organisms from exposure to mixtures of many contaminating compounds can be very different from the mere sum of the effect of individual compounds (Backhaus et al. 2008).

It is obvious that whole effluent toxicity data on scrubber water from ships is much needed in order to make sound assessments of the possible adverse effects the discharged water may have on marine ecosystems. In the EMERGE project researchers from universities and research institutes in Sweden, Great Britain, Portugal, Italy, and Greece have carried out a range of toxicity tests of scrubber water on marine organisms, and their different life stages, and this report presents a compilation of the obtained data. All testing was done on whole scrubber water, either from ships in operation, or from a pilot scrubber equipment. The water was collected and treated so that it would be as similar as possible to the discharged scrubber water that marine organisms would encounter in the field. Test organisms included both single marine species, and here several different species and different life stages of the same species were used, and natural communities of species.

3 METHODOLOGY

Scrubber water was either collected from open loop systems onboard ships in operation, or from a pilot system at Chalmers University of Technology (Gothenburg, Sweden) where a scrubber unit in stainless steel was connected to a four-cylinder 100 kW engine from Volvo Penta. The fuels used on board the ships were high sulphur HFO, with a sulphur content of approximately 2.5%. In the Chalmers pilot unit was used Heavy Gas Oil (HGO), with a considerably lower sulphur content, 0.3 – 0.7%. The scrubber water collected from ships in operation mainly originated from the ship Leo C (DANAOS Shipping co. LTD). Scrubber water was sampled at thirteen sites along the ship´s route between the departure port in Belgium and the arrival port in Turkey during December 2021 (Figure 3-1). During its route the ship was using different fuels and combination of fuels depending on the restrictions of the particular sea area. The ship was also running on different engine powers at the different times of sampling. The water was collected as the ships passed through different sea areas i.e., the English Channel, the North Sea, the North-East Atlantic and the Mediterranean (Figure 3-1). During the Covid-19 lockdown, some EMERGE-partners had the opportunity to get scrubber water from another ship, Catherine C (DANAOS Shipping co. LTD) and were also allowed to run experiments at their respective institutes. Ecotoxicological results are thus available from these studies, and they are presented here. Unfortunately, chemical analyses were not performed on all these waters and are thus presented as Not Available (N/A) in this report. The time of sampling is unfortunately unclear.

The tested waters were hence not identical in their chemical composition, but this reflects real life situations under which scrubber water is produced. Sampling on board the ship was done as close to the discharge point to the sea as possible in order to obtain a water similar to that which the animals in the field are exposed to. Aside from the varying contents of pollutants, the scrubber waters also differ in pH when released into the sea. IMO regulations allow environmental deposition of scrubber water provided that with a pH is no less than 6.5 measured at the ship's overboard discharge (IMO 2015). Scrubber water has been recognized for contributing to ocean acidification locally (Hassellöv et al. 2013) and reduced sea water pH is known to affect a range of marine organisms (Thor and Dupont 2018). The pH of the scrubber water was therefore not manipulated in the ecotoxicological tests of the EMERGE project other than being naturally increased by dilution with seawater.

The water was collected in acid (HCl) and acetone washed glass bottles, kept cool (+4 – 8°C) and dark and was transported to the laboratories within a week. Subsamples of the scrubber waters were

sent to the Catalan Institute for Water Research (ICRA) for chemical characterization of metal, PAH and alkylated PAH contents. A comprehensive presentation of the analytical results can be found in the D2.2 report of the EMERGE project.

Figure 3-1. Sampling sites for scrubber water from the ship Leo C (DANAOS Shipping co. LTD) on the route from Belgium to Turkey. White points with black outline indicate scrubber water used in ecotoxicological tests and experiments. Sites 1: English Channel early, 2: English Channel, 3: North Atlantic outside SECA, 4: Atlantic Bay of Biscay, 5: Atlantic North Portugal, 6: Before Gibraltar Atlantic side, 7: Gibraltar, 8: After Gibraltar, 9: Mediterranean, 10: Mediterranean South of Italy, 11: Mediterranean Greece, 12: Athens, 13: Aegean Sea.

The different research groups received scrubber water sampled at different times, in different sea areas and sometimes also from different ships, meaning that the chemical composition of the tested waters was not identical. The original intention was that the scrubber water tested should derive from ships as they passed through waters close to the EMERGE case-study areas (Figure 3-2) place where the experiments were conducted with organisms endemic to the particular area, but due to logistic issues this was not always possible.

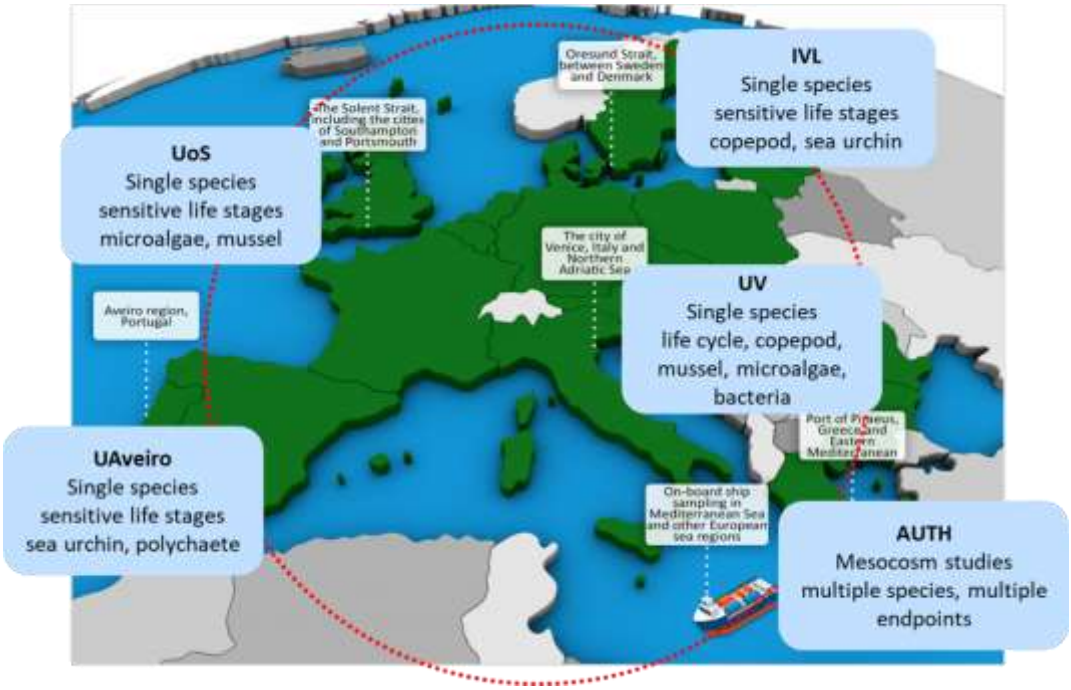


Figure 3-2. Universities/research institutes where ecotoxicological tests of scrubber water were carried out and the types of tests conducted. UoS: University of Southampton, Great Britain; UAV: University of Aveiro, Portugal; UV: University of Venice, Italy; AUTH: Aristotle University of Thessaloniki, Greece; IVL: IVL Swedish Environmental Research Institute, Sweden. Map also indicates the Case study areas and the On-board campaign.

Ecotoxicological tests were carried out at five research laboratories at University of Venice (UV) in Italy, University of Southampton (UoS) in Great Britain, IVL Swedish Environmental Research Institute (IVL) in Sweden, University of Aveiro (UAV) in Portugal, and Aristotle University of Thessaloniki (AUTH) in Greece (Figure 3-2). At all laboratories but AUTH, experiments were carried out with just one species at the time. At AUTH natural plankton communities, with a range of different microorganisms, were collected and exposed to the scrubber water in mesocosm exposures.

All organisms selected for the tests in the EMERGE project are pelagic or have pelagic larvae, that is, they spend their entire life, or part of their life cycle, in the water column. This choice was driven by the consideration that the scrubber water is discharged close to the water surface and the organisms of the pelagic ecosystem are hence those that will be exposed to the highest concentrations of the hazardous compounds.

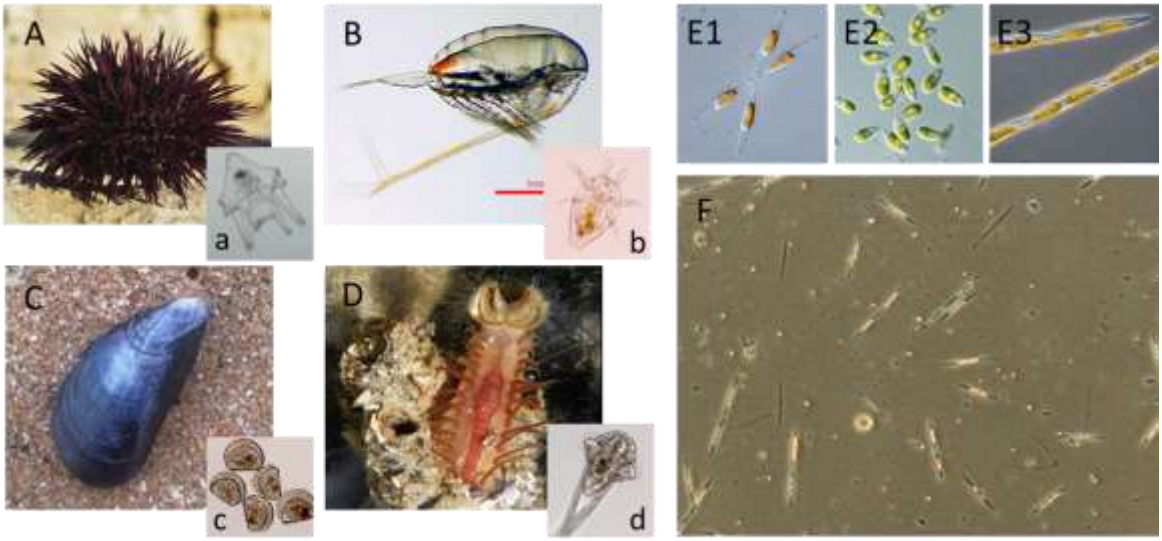


Figure 3-3. Images of some species, some of their larvae and planktonic communities used in the ecotoxicological experiments and tests with scrubber water. A: Sea urchin (*Paracentrotus lividus*, Echinodermata) with pluteus larva (a), B: Copepod (*Calanus sp*., Crustacea) with nauplius larva (b), C: Blue mussel (*Mytilus sp*., Mollusca) with veliger larva (c), D: Polychaete (*Sabellaria alveolate*, Annelida) with trochophore larva (d), E: Microalgae of the species *Phaeodactylum tricornutum* (E1), *Dunaliella tertiolecta* (E2) and *Pseudo-nitzschia sp.* (E3), F: Phytoplankton community dominated by the microalga *Pseudo-nitzschia cf. pungens* (live and dead – empty cells) in 10% scrubber water (station 11). A broken cell of *IU* and combustion particles are also shown, Micrograph by Moustaka M.

The tested species in the single species tests carried out at UV, UoS, IVL and UAV were either marine algae, marine invertebrates, or marine bacteria. Of these, the truly pelagic ones are the copepods *Acartia tonsa* and *Calanus helgolandicus*, and the phytoplankton species *Phaeodactylum tricornutum* and *Dunaliella tertiolecta*. Other test species spend their adult life on the sea floor whereas the life stages tested in the present project, eggs and larvae, are planktonic. These are the sea urchins *Strongylocentrotus droebachiensis* and *Paracentrotus lividus*, the polychaete *Sabellaria alveolata* and the blue mussels *Mytilus galloprovincialis* and *Mytilus edulis*. The mesocosm experiments conducted by AUTH on natural field collected communities included organisms <200 μm, mainly numerous species of phyto- and bacterioplankton that were present at the time of the sampling. Pictures of the organisms and communities are shown in Figure 3-3.

In most single species experiments organisms were exposed to scrubber water in a range of five concentrations between 0.001 and 40%, expressed as % of scrubber water in natural sea water (volume/volume). The lowest concentration was adjusted twice during the project, since effects were observed already at the lowest concentration tested. Thus, all tests do not have the same lowest exposure concentration, the lowest concentration tested within the project by some groups was

0.0001% scrubber water. From the generated data the lowest concentration found to have an effect that was statistically different from control organisms exposed to sea water only (Lowest Observed Effect Concentration, LOEC), and the highest concentration with no statistically significant effect (No Observed Effect Concentration, NOEC) were calculated. The response data was also used to estimate the Effective Concentration or Lethal Concentration 10% (EC10/LC10), that is, the scrubber water concentration having a 10% effect on the analysed endpoint or causing the death of 10% of the exposed population of test organisms, and EC50% or LC50% (EC50/LC50), the scrubber water concentration having a 50% effect on the analysed endpoint or causing the death of 50% of the exposed population of test organism. EC/LC10 and EC/LC50 were estimated from a dose-response curve based on the obtained data on effect/mortality among scrubber water exposed organisms compared to a control group exposed to sea water only. LOEC and NOEC are hence actual concentrations to which the organisms were exposed, whereas EC10 and EC50 are statistical estimates calculated from a curve fit to the obtained response data.

In the mesocosm experiments, communities of microplankton (Figure 3-3F) were collected in two polluted (a port and a marina) and two unpolluted sites, and the organisms were exposed to either 1 and 10%, or 1, 2 and 5% scrubber water. The total abundance (number of cells) was calculated for three groups of the plankton, microphytoplankton (> 2 µm), picophytoplankton (approximately 0.5 - 2 µm) and bacterioplankton. Analysis was also done on population density of a selected number of individual species.

4 RESULTS

4.1 Toxicity of scrubber water – results from ecotoxicological experiments

In this section a summary of the results from all ecotoxicological experiments on scrubber water carried out within the EMERGE project is presented. More detailed information on experimental set-ups and obtained data is found in the appendices where each research group presents the work they have carried out:

- University of Venice: Appendix 1
- University of Southampton: Appendix 2
- IVL Swedish Environmental Research Institute: Appendix 3
- University of Aveiro: Appendix 4
- Aristotle University of Thessaloniki: Appendix 5.

The data indicate that invertebrates are more sensitive to scrubber water than phytoplankton, it also shows that some of the early invertebrate life stages are much more sensitive than others. It should however be noted that "phytoplankton" is a very large and diverse group, and the sensitivity between groups of phytoplankton species may differ. Table 4-1 summarises the most sensitive endpoints and their LOECs and NOECs in response to the percentage of scrubber water mixed into the seawater the animals were exposed to. In some cases, experiments with the same design have been repeated with different scrubber waters which has resulted in different LOECs. This is not surprising, not least because the chemical composition of the scrubber water can be very different depending on a range of factors, e.g., what fuel was used, the condition of the specific ship engine and the engine power when the scrubber water was produced (Lunde Hermansson et al. 2021). However, variations in scrubber water composition depending on what ship is passing is a reality also for organisms in the field. In Table 4-1 we follow the precautionary principle and present the lowest scrubber water concentrations found to have a statistically significant effect on the most sensitive endpoints. However, a more extended compilation of data from all experiments is presented in Tables 4-2 – 4-5, and more detailed descriptions of the experiments are presented by each research group in appendices 1-5. In many experiments LOEC was equal to the lowest scrubber water concentration tested. In these cases, the no observed effect concentration, NOEC, had to be expressed as <LOEC, i.e., an unknown concentration that is lower than LOEC.

The most sensitive endpoint in response to scrubber water detected in the EMERGE project was the fertilisation success of eggs from the green sea urchin *Strongylocentrotus droebachiensis*,

where the LOEC was found at the lowest tested concentration, 0.0001% scrubber water (1 mL scrubber water per m^3 water) (Table 4-1). This was, however, the only laboratory where such a low concentration was tested. The second most sensitive endpoints were malformation of the larvae of the sea urchin *Paracentrotus lividus*, the polychaete *Sabella alveolata* and the blue mussel *Mytilus edulis* with LOECs of 0.001% scrubber water. Even in these tests LOEC was equal to the lowest tested scrubber water concentration. At 0.01% scrubber water also the fertilization success of sea urchin eggs and the egg production of the copepod *Acartia tonsa* were significantly reduced.

Table 4–1. Summary of the most sensitive endpoints and the lowest concentrations of scrubber water causing a statistically significant toxic response. Concentrations are expressed as percentage of scrubber water in the exposure water. NOEC (No Observed Effect Concentration) = the highest tested concentration of scrubber water with no significant effect, LOEC (Lowest Observed Effect Concentration) = the lowest tested concentration of scrubber water having a significant effect.

End point	Species	NOEC (%)	LOEC (%)
Copepod, egg production[1, A]	*Acartia tonsa*	0.001	0.01
Copepod, larval development[1, A]	*Acartia tonsa*	<0.01	0.01
Sea urchin, egg fertilization[2, B]	*Strongylocentrotus droebachiensis*	<0.0001	0.0001
Sea urchin, malformation of larvae[2, B]	*Strongylocentrotus droebachiensis*	0.01	0.1
Sea urchin, egg fertilization[3, A & C]	*Paracentrotus lividus*	<0.01	0.01
Sea urchin, malformation of larvae[3, C]	*Paracentrotus lividus*	<0.001	0.001
Polychaete, malformation of larvae[3, B]	*Sabellaria alveolata*	<0.001	0.001
Blue mussel, malformation of larvae[4, D]	*Mytilus edulis*	<0.001	0.001

[1]UV, [2]IVL, [3]UAV, [4]UoS, [A]tests were done with Chalmers scrubber water, [B]tests were done with LeoC 1 scrubber water, [C]tests were done with Catherine C scrubber water, [D]tests were done with Calmers and CatherineC N/A scrubber water.

Mortality is a common endpoint in standardised toxicity tests. In EMERGE increased mortality in response to scrubber water exposure was tested on different life stages of two species of copepods, adult individuals and nauplii larvae of *Acartia tonsa* and juveniles of *Calanus helgolandicus*. *Acartia* adults and larvae were found to be relatively insensitive, with NOECs being 10 and 5%, respectively, and LOEC 20 and 10%, respectively (Tables 4-2A & 4-3), whereas for juvenile Calanus the lowest tested concentration, which was 1%, caused a significantly increased mortality compared to control animals (Table 4-3). The difference in sensitivity could be caused by a difference between species, between life stages or it could be connected to the fact that the *Acartia* experiments were carried out on animals reared in the laboratory, whereas the *Calanus* were field collected a few days before the experiments. The difference can also be attributed to the fact that different scrubber waters were used.

Table 4–2A & B. Toxicity of scrubber water on marine invertebrates. Concentrations are expressed as percentage of scrubber water in the exposure water. NOEC (No Observed Effect Concentration); LOEC (Low Effect Concentration); EC10/EC50 = scrubber water concentration having a 10/50% inhibiting effect on the tested endpoint.

A. Species	Endpoint	NOEC (%)	LOEC (%)	EC10 (%)	EC50 (%)
Acartia tonsa, copepod[1]	Adults, mortality	5	10	-	-
	Egg production, (LeoC 10, exp 1)	0.001	0.01	-	-
	Egg hatching, (LeoC 10, exp 1)	>1	>1	-	-
	Larval survival, (LeoC 10, exp 1)	>1	>1	-	-
	Larval development, (LeoC10, exp 1)	0.01	0.1	-	-

	Egg hatching, (LeoC 10, exp 2)	10	20	-	-
	Larval survival, (LeoC 10, exp 2)	10	20	-	-
	Larval development, (LeoC 10, exp 2)	1	2	-	-
Calanus helgolandicus, copepod[2, A]	Copepodite mortality	<1	1	-	-
	Copepodite moulting	<1	1	-	-
	Larvae, malformation	0.01	0.1	2.7	4.7

[1]UV, [2]IVL, [A]Experiments were done with scrubber water defined in Thor et al. (2021).

B. Species	Endpoint	NOEC (%)	LOEC (%)	EC10 (%)	EC50 (%)
Strongylocentratus droebachiensis, sea urchin[2]	Egg, fertilization success (LeoC 1)	<0.0001	0.0001	2.3	7.7
	Larvae, malformation (LeoC 1)	0.01	0.1	2.7	4.7
Paracentrotus lividus, sea urchin[3]	Egg, fertilization success, (CatherineC, exp. 2)	<0.01	0.01	7.6	33.7
	Larvae, malformation, (CatherineC, exp 2)	<0.001	0.001	-	6.13
	Egg, fertilization success, (LeoC 1)	0.01	0.1	7.2	11.4
	Larvae, malformation, (LeoC 1)	0.1	0.01	0.78	5.5
	Egg, fertilization success, (CatherineC, exp. 1)	1.56	3.13	-	22.9
	Larvae, malformation, (CatherineC exp. 1)	<1.56	1.56	-	1.5
Sabellaria alveolate, polychaete[3]	Larvae, malformation (CatherineC)	0.001	0.01	-	9.44
	Larvae, malformation (LeoC 1	<0.001	0.001	1.13	10.5

Mytilus galloprovincialis, mussel[1]	Larval development (LeoC 10)	0.1	1	4.9	6
Mytilus edulis, mussel[4]	Larvae, malformation (CatherineC N/A)*	<0.001	0.001	0.27	9.27
Mytilus edulis, mussel[4]	Larvae, malformation (Chalmers)	<0.001	0.001	0.06	0.54

[1]UV, [2]IVL, [3]UAV, [4]UoS. *CatherineC N/A refers to a scrubber water collected onboard the ship CatherineC during the EMERGE project, but the water was never analyzed chemically.

Table 4–3. Ecotoxicological tests of scrubber water from the pilot scale scrubber equipment (Chalmers) on marine invertebrates. Concentrations are expressed as percentage of scrubber water in the exposure water.

Species	Endpoint	NOEC (%)	LOEC (%)	EC10 (%)	EC50 (%)
Acartia tonsa, copopod[1]	Adult, mortality	20	40	-	-
Acartia tonsa[1]	Adult females, egg production	<0.01	0.01	-	-
	Egg hatching	20	40	-	-
	Nauplia stage VI, survival	10	20	-	-
	Larval development, egg to copepodite	<0.01	0.01	-	-
Paracentrotus lividus, sea urchin[2]	Egg fertilization (experiment 1)	<0.01	0.01	6.36	26.68
	Larvae, malformation (experiment 1)	0.001	0.01	0.265	8.04
	Egg fertilization (experiment 2)	1.56	3.125	-	13.7
	Larvae, malformation (experiment 2)	<1.56	1.56	-	1.3

Sabellaria alveolate, polychaete[2]	Larvae, malformation	0.001	0.01	<0.001	3.8
Aliivibrio fisheri, bacteria[1]	Bioluminescence, inhibition	10	20	24.8	-

[1]UV, [2]UAV.

The impact of scrubber water on microalgae in both single species tests (Table 4-4) and in studies of natural populations (Table 4-5) was, in the performed experiments, found to be around 20,000 to 40,000 times (>four orders of magnitude) less sensitive than the development of invertebrate larvae (Table 4-1). The LOEC for effects on the growth rate of two species of phytoplankton, *Phaeodactylum tricornutum* and *Dunaliella tertiolecta*, was found to be 20 and 40% scrubber water, respectively, compared to the LOECs of 0.0001 and 0.001% for the most sensitive early life stages of marine invertebrates (Tables 4-1─ 4-3). Data on the effect of scrubber water on population density of natural multispecies plankton communities showed a toxic effect on the microphytoplankton (>2 µm) to the highest test concentrations, 10 or 5% scrubber water, but not in the second lowest concentrations, which were 1 or 2% (Table 4-5). Picoplankton (0.5 - 2 µm) were only analysed in two of the communities and in one of them a negative effect was detected at exposure to 10% scrubber water whereas no toxic effect could be observed in the other. The bacterioplankton community was not affected even at the highest scrubber water concentration at any of the sites. Still, the collected data from the experiments on phytoplankton exposed to scrubber water, both for single species and communities, strongly indicate that they are less sensitive than early invertebrate development.

Table 4–4. Toxicity of scrubber water on planktonic microalgae. Concentrations are expressed as percentage of scrubber water in the exposure water. NOEC (No Observed Effect Concentration); LOEC (Low Effect Concentration); EC10/EC50 = scrubber water concentration having a 10/50% inhibiting effect on the tested endpoint. All tests were done by UV with LEOC 10 scrubber water.

Species	Endpoint	NOEC (%)	LOEC (%)	EC10 (%)	EC50 (%)
Phaeodactylum tricornutum, algae	Growth rate	20	40	34	-
Dunaliella tertiolecta, algae	Growth rate	10	20	15	-

Aliivibrio fisheri, bacteria	Bioluminescence, inhibition	10	20	23	39

Table 4–5. Toxicity tests of the effect of scrubber water on natural communities of planktonic microorganisms from four sites. Total abundance was analysed for the microorganisms divided into microphytoplankton, picophytoplankton and bacterioplankton. Highest test concentration for community 1 and 2 was 10%, and for 3 and 4, 5%. Exposure concentrations are set as % scrubber water. All tests were performed by the AUTH on LeoC 11 scrubber water.

End point	Species	Site	NOEC (%)	LOEC (%)
Total abundance	Microphytoplankton community 1	Thessaloniki port	1	10
Total abundance	Picophytoplankton, community 1	Thessaloniki port	1	10
Total abundance	Bacterioplankton, community 1	Thessaloniki port		>10
Total abundance	Microphytoplankton community 2	Plagia, "unpolluted" site	1	10
Total abundance	Picophytoplankton, community 2	Plagia, "unpolluted" site		>10
Total abundance	Bacterioplankton, community 2	Plagia, "unpolluted" site		>10
Total abundance	Microphytoplankton community 3	Flisvos Marina	2	5
Total abundance	Bacterioplankton, community 3	Flisvos Marina		>5
Total abundance	Microphytoplankton community 4	Vouliagmeni, "unpolluted" site	2	5
Total abundance	Bacterioplankton, community 4	Vouliagmeni, "unpolluted" site		>5

Tests were also carried out on a single marine bacterium of the species *Aliivibrio fisheri.* This is a standardized test where the endpoint is the inhibition of the bacterial bioluminescence, and it is frequently used to study toxicity in wastewater. The bioluminescence is part of the respiratory process of the bacteria, and an inhibition is hence a sign of a disturbance on a basic metabolic level. Just as for the natural communities of bacterioplankton the sensitivity to scrubber water was very low (Tables 4-2, 4-3 and 4-5), with LOEC and NOEC being 40 and 20%, respectively.

The toxicity of the scrubber water from the pilot scale scrubber equipment was in a similar range as the real scrubber water for most parameters (Table 4-3). In some cases, it was less toxic e.g., a significant malformation of larvae of *Sabellaria alveolate* was occurring at a ten percent higher

concentration, 0.01% in pilot scrubber water compared to 0.001% in real scrubber water collected on a ship.

It is beyond the scope of this study to indicate which contaminants were responsible for the observed toxicological effects. It is nevertheless interesting to see what the concentrations of some of the potentially most toxic chemicals in scrubber water were when statistically significant effects were detected. A significant effect on fertilisation of sea urchin eggs occurred when the concentration of 16 PAH was 0.02 ng/L of exposure water, alkylated PAHs 0.04 ng/L, vanadium 0.4 ng/L, zinc 0.05 ng/L and copper 0.01 ng/L. Malformation of larvae of sea urchin, polychaete and blue occurred at scrubber water concentrations of 0.001%, which corresponded to approximately 0.08 ng 16 PAH/L, 0.13 ng alkylated PAH/L and 1.0, 0.4 and 0.07 ng/L for the metals vanadium, zinc and copper. These concentrations are far below various threshold levels set by national and international agencies, an issue that will be discussed in more detail in section 5.2 Toxic effects cannot be determined based on chemical analysis alone. It should be noted that the chemical composition of the scrubber water used was not identical in all experiments and concentrations of 16 PAHs, alkylated PAHs, and metals of the scrubber waters used by the different groups is presented in Tables 4-6 to 4-8.

There are other factors but toxic substances that could add to the observed effects. Reduced pH caused by the acidic scrubber water is one of them. However, in the lowest concentrations with a significant toxic effect, 0.0001 – 0.001% scrubber water in the exposure water, the change in pH was negligible. It is also noteworthy that the toxicity of the scrubber water from the pilot scale equipment, for many endpoints, was in the same range as the scrubber water from ships in operation, even though the former used a fuel with considerably lower sulphur content than the latter, 0.3 – 0.7% compared to ~2.5% sulphur. The various kinds of combustion particles present in the scrubber water can also be expected to have contributed to the toxicity. Many of these particles have high concentrations of PACs and metal adsorbed to the surface, but it is still not clear how available these contaminants are for uptake in living organisms.

4.2 Data on the chemical composition and similarities of scrubber waters

The similarity of all scrubber water samples i.e., 12 samples scrubber waters collected during the DANAOS Leo C cruise from Belgium to Turkey (scrubber water from LeoC sites 1-11 & 13) (Figure 3-1), samples from the ship Catherine C and scrubber water produced in the pilot device at Chalmers University of Technology and the five scrubber waters used in ecotoxicological tests

(scrubber water from LeoC sites 1, 10 and 11, Chalmers and CatherineC used by UAV) was investigated using multivariate statistics. Multivariate analyses were performed using the PRIMER-E (6.1.13) software (Clarke and Gorley 2006) by applying a Bray-Curtis similarity on square root transformed abundance data on individual chemical compounds in unfiltered samples. Due to the lack of replicate samples for each water, similarity percentages are based on clustering calculated using SIMPROF (PRIMER-E (6.1.13)). Statistical analyses were made for the compounds included in the chemical analyses of the 16 US EPA PAHs, alkylated PAHs and metals separately for all 12 scrubber water samples together and for the five scrubber waters used for ecotoxicological testing alone i.e., those originating from Chalmers (used by UAV and UV), CatherineC (used by UAV), LeoC 1 (used by IVL and UAV), LeoC 10 (used by UV) and LeoC 11 (used by AUTH) (Figure 4-1).

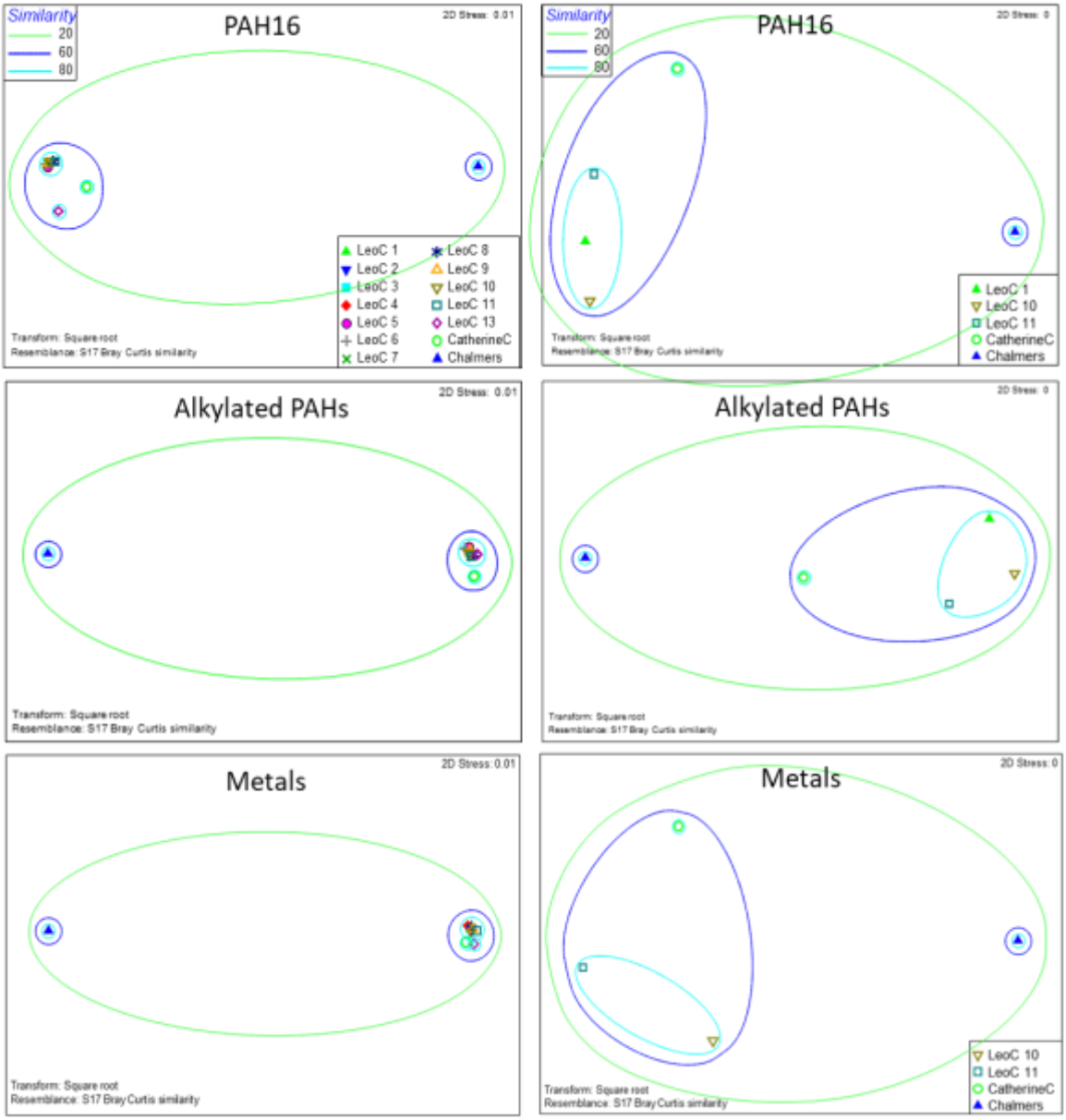


Figure 4-1. Non-metric multi-dimensional scaling (nMDS) projections of chemical compositions of all original scrubber waters analysed (left column) and scrubber waters used for ecotoxicological testing (right column). Analyses are based on Bray-Curtis similarity on square root transformed abundance data of individual chemical compounds in unfiltered samples. PAH16: the 16 US EPA PAHs are included, Alkylated PAHs: the measured alkylated PAHs are included, Metals: V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Cd, Hg, Pb, U are included. Degrees of similarity are shown as coloured ovals indicating 20, 60 and 80% similarity.

From the non-metric multidimensional scaling (nMDS) clustering it is apparent that the Chalmers scrubber water differs substantially in composition from all the other scrubber waters. All other

scrubber waters are only 20% similar to the Chalmer scrubber water regarding 16 PAHs, alkylated PAHs and metals individually. Furthermore, the LeoC scrubber waters (1-13) are more similar to each other (80%) than to the CatherineC scrubber water (60%) considering all three contaminant groups separately.

Concentrations of individual compounds in the five scrubber waters used in ecotoxicological tests are presented in Tables 4-6 to 4-8. For a detailed description of the chemical contents of all scrubber analysed waters within EMERGE see D2.2. The concentrations of 16 US EPA PAHs and alkylated PAHs were significantly lower in the Chalmers scrubber water compared to the other scrubber waters used in ecotoxicological experiments and tests (Tables 4-6 & 4-7). The sum of 16 US EPA PAHs and of alkylated PAHs for the CatherineC scrubber water were approximately half of that of the LeoC scrubber waters (Tables 4-6 & 4-7). The Chalmers scrubber water contained significantly higher concentrations of iron and copper compared to the other three scrubber waters used in the ecotoxicological tests (Table 4-8). Chrome and manganese concentrations were also higher while vanadium and uranium concentrations were comparably lower. Concentrations of vanadium, nickel and iron were lower in the CatherineC scrubber water compared to the LeoC scrubber waters.

Table 4–6. Concentrations of individual compounds and the total sum (ng/L) of the 16 US EPA PAHs in scrubber water used in the ecotoxicological tests, data from unfiltered samples.

Compound	LeoC 1[2,3] (ng/L)	LeoC 10[1] (ng/L)	LeoC 11[4] (ng/L)	CatherineC[3] (ng/L)	Chalmers[1,3] (ng/L)
Naphthalene	8958	7499	6532	2371	267
Acenaphthylene	67.6	141	144	34.8	<LOD (0.33)
Acenaphthene	387	592	284	52.6	<LOD (0.33)
Fluorene	1135	1000	1378	934	31.4
Phenanthrene	4499	4815	3460	3854	223
Anthracene	<LOD (1.67)	<LOQ (5.00)	<LOQ (5.00)	<LOD (0.33)	<LOD (0.33)
Fluoranthene	195	139	135	128	23.2
Pyrene	475	701	335	167	44.6
Benzo(a)anthracene	18.4	89.1	16.3	<LOD (0.33)	<LOD (0.33)
Chrysene	214	236	187	62.6	10.6
Benzo(b)fluoranthene	26.7	15.1	7.9	6.56	<LOD (1.67)

Benzo(k)fluoranthene	9.2	<LOQ (2.51)	<LOQ (2.51)	<LOD (1.67)	<LOD (1.67)
Benzo(a)pyrene	<LOD (3.70)	<LOD (3.70)	<LOD (3.70)	<LOD (1.67)	<LOD (1.67)
Indeno(1,2,3-cd)pyrene	<LOD (1.67)	<LOD (1.67)	<LOD (1.67)	<LOD (1.67)	<LOD (1.67)
Dibenzo(a,h)anthracen	29.4	55.9	9.9	<LOD (1.67)	<LOD (1.67)
Benzo(g,h,i)perylene	175	111	39.0	<LOD (1.67)	<LOD (1.67)
Sum 16 PAH	16188	15394	12527	7611	600

Universities/institutes using the different scrubber waters: [1]UV, [2]IVL, [3]UAV, [4]AUTH

Table 4–7. Concentrations of individual compounds and the total sum (ng/L) of alkylated PAHs in scrubber water used in the ecotoxicological tests, data from unfiltered samples.

Compound	LeoC 1[2,3] (ng/L)	LeoC 10[1] (ng/L)	LeoC 11[4] (ng/L)	CatherineC[3] (ng/L)	Chalmers[1,3] (ng/L)
2-methylnaphthalene	4969	4599	2806	1005	37.0
1-methylnaphthalene	4521	3573	2437	817	56.3
C2-methylnaphthalene	7541	8461	5839	1045	54.4
C3-methylnaphthalene	4426	4025	3286	422	<LOD (0.33)
C4-methylnaphthalene	1407	1334	1354	195	<LOD (0.33)
C1-methylphenanthrene	4644	4961	4685	4363	323
C2-methylphenanthrene	2869	3405	3001	2138	238
C3-metylphenanthrene	1089	1484	1224	460	57.9
C4-methylphenanthrene	640	921	703	250	<LOD (0.33)
C1-methylfluorene	2139	1859	1559	1431	64.5
C2-methylfluorene	1243	1101	986	850	76.7
C1-methylfluoranthene/pyrene	629	889	484	203	33.0
Sum alkylated PAH	36118	36612	28364	13178	940

Universities/institutes using the different scrubber waters: [1]UV, [2]IVL, [3]UAV, [4]AUTH

Table 4–8. Concentrations of single metal elements (µg/L) in scrubber water used in the ecotoxicological tests. Sample vials of scrubber water from LeoC 1 were broken and could not be analysed for metals.

Element	LeoC 10[1] (µg/L)	LeoC 11[4] (µg/L)	CatherineC[3] (µg/L)	Chalmers[1,3] (µg/L)
Vanadium	348.3	226.7	96.8	1.9
Chromium	7.5	4.56	6.71	65.4
Manganese	3.08	0.89	1.57	65.2
Iron	109.2	64.1	47.8	570.0
Cobalt	0.4737	0.43	0.334	1.3
Nickel	99.5	72.8	32.9	90.2
Copper	5.33	1.41	6.97	87.7
Zinc	48.8	25.7	41.7	34.8
Arsenic	2.02	1.94	1.87	1.2
Cadmium	0.0316	0.0171	0.0347	0.0
Mercury	0.0018	0.0012	0.0016	0.0
Lead	0.3516	0.0902	0.921	1.7
Uranium	6.89	6.98	3.59	0.8

Universities/institutes using the different scrubber waters: [1]UV, [2]IVL, [3]UAV, [4]AUTH

4.3 Evaluation of EMERGE ecotoxicological experiments and tests using the Criteria for Reporting and Evaluating Ecotoxicity Data (CRED)

The ecotoxicological work carried out in the EMERGE project is a mixture of experiments run according to standardized protocols, modified standardized protocols and those designed by the scientists conducting them. All ecotoxicological scientists are fundamentally concerned with the quality, reliability and relevance of their work. Ecological relevance, i.e., the applicability of the results in real world exposure situations, is a very important aspect of ecotoxicology and influence the utility of the generated data for risk/impact assessment. If the three abovementioned criteria are not met the work will not pass the rigorous peer-review process of scientific journals and the work will thus not fit for publication. So, although ecotoxicological data published in technical reports may be said to follow standardized protocols, it will not guarantee that the quality of the data is sufficiently high for scientific journals or fit to serve as a sound basis for regulatory decisions. There are several examples of reports on scrubber water toxicity that have received significant

attention and have had a major influence on scrubber water legislation, but where the results could be questioned due to how the scrubber water was handled prior to the experiments and how the experiments were conducted.

In order to assure the quality and usefulness in relation to risk assessment (development of Environmental Quality Standards-EQS) of the ecotoxicological data produced within the EMERGE project, the EMERGE-Ecotoxicology group has chosen to follow the technical guidance document for deriving environmental quality standards issued by the European Commission (TGD-27) (EC 2018). The TGD-27 states that:

"*Studies do not need to have been performed under a formal quality assurance scheme, such as Good Laboratory Practice (GLP) or do not need to be OECD validated or ISO certified, but should follow generally accepted good scientific principles*."

and that:

"*Studies that might influence an EQS must be quality assessed. The assessment may be performed according to the scheme developed by (Klimisch et al. 1997) or CRED (Kase et al. 2016, Moermond et al. 2016). The Klimisch system is a long-established one that is also used in other chemical assessment regimes, but CRED offers the ability to further assess relevance of aquatic ecotoxicity data in addition to the reliability criteria and is recommended to be applied for the critical studies in a dataset.*"

The EMERGE-Ecotoxicology group has chosen to use the CRED method (Moermond et al. 2016) to evaluate the transparency, harmonisation and quality assurance of the ecotoxicological experiments and data produced within the EMERGE project, in view of their use and evaluation for risk/impact assessment. The CRED method uses a set of 20 **reliability** and 13 **relevance** criteria, which support the harmonization in the evaluation and reporting of ecotoxicological results (inherently subject to expert judgement)

Each research team assessed its own ecotoxicological experiments and results against CRED criteria. The CRED reports from the individual laboratories (universities/research institutes) for individual and sets of studies are attached as appendices 6 - 11.

5 DISCUSSION

5.1 Discussion of the results of ecotoxicological testing within EMERGE

The experimental data generated in the EMERGE project reveals that the scrubber water has a negative effect on marine organisms at much lower concentrations than previously known. The most sensitive endpoints in the most sensitive species were significantly affected already with 0.0001 - 0.001% scrubber water in the exposure water (Table 4-1). Of vital ecological importance is the high sensitivity found in some of the early life stages of the pelagic copepods *Acartia tonsa* and *Calanus helgolandicus*. These small crustaceans are important grazers on microalgae and are themselves essential prey for pelagic fish, such as cod. The cod populations around the world have declined drastically over the past decades, and although over-fishing is an important reason for this, there is convincing evidence in the scientific literature that the reduction in cod also is caused by a reduction in populations of pelagic copepods (Beaugrand et al. 2003, Heath and Lough 2007, Beaugrand and Kirby 2010a). The reason for declining copepod populations is generally suggested to be climate change, and the role of toxicants is rarely considered as a relevant factor. However, the new data from the EMERGE experiments on how surprisingly sensitive certain copepod life stages are to the pollutants in the scrubber water may contribute to draw a more complex picture showing that also pollutants could play a part in the decline of copepod populations.

The observed effects on the free-swimming larvae of bottom living organisms at very low concentrations could have a serious impact on both soft bottom and hard bottom marine ecosystems. There has been a dramatic decline in some bottom-living species over the past year, perhaps one of the more noticeable is the blue mussel in the North Atlantic where the populations have gone from very abundant to rare in just over a decade (Baden et al. 2021). Again, there are most likely several reasons for this, but that pollution is no one of them is far from excluded.

A strength in the ecotoxicological investigations of EMERGE has been the range of different experiments on scrubber water toxicity that has been carried out. Different species and different life stages and end points of the species have been tested and been found to respond very differently to the scrubber water exposure. Some endpoints included in the EMERGE experiments were found to be relatively insensitive to scrubber water exposure. This was the case for, e.g., growth rate of the algal species *Phaeodactylum tricornutum* and *Dunaliella tertiolecta*, and mortality of adult individuals and larvae of the copepod species *Acartia tonsa* (Tables 4-2 – 4-4). Although increased mortality was found to be a sensitive endpoint in juveniles of another copepod species, *Calanus helgolandicus*, the data from the EMERGE experiments indicate that neither algae growth rate nor

mortality of small crustaceans can be considered as suitable endpoints to assess the toxicity of scrubber. These endpoints, growth rates of microalgae and mortality of small crustaceans, are included in many standardized protocols and have been used also in toxicity studies of scrubber water, e.g., in a Japanese study that has received much attention (MEPC 74/INF.24, 2018). In that study the ecotoxicity testing of whole scrubber water was carried out on microalgae growth rate, on mortality of specimens of the crustacean *Hyale barbicornis* and on fish of the species *Oryzias javanicus* (MEPC-IMO 2018). The EC50 and LC50 were high for all parameters (this means that the toxicity was low); EC50 for the algae growth inhibition was estimated to around 50% scrubber water in the exposures water, LC50 was around 20% scrubber water for the crustaceans and around 35% for fish. The conclusion drawn from these results was that "there is no effect [*of scrubber water*] on all marine organisms". If the EMERGE experiments had been limited to study just algae growth and mortality on small crustaceans (like *Acartia tonsa*), the conclusions could have been the same. However, by widening the scope and including several different life stages of organisms a more realistic picture of the toxicity of the scrubber water has emerged.

5.2 Toxic effects cannot be determined based on chemical analyses alone

The acceptance to allow the use of scrubbers as a means of complying with the sulphur cap instead of switching to cleaner fuels came without any kind of risk assessment of what would be the effect on the marine ecosystems. According to DNV GL the first scrubbers for exhaust gases from ships were installed in 2007, in 2015 there were 242 ships, and in 2022 4737 ships globally (the latter figure includes scrubbers on order). Around 85% of them have an open loop system. Each of these ships discharge scrubber water continuously at a rate that varies that often may be in the order of 1000 m^3/hour as long as the engine is running (Lunde Hermansson et al. 2021). Still, this large new source of pollutants to the seas has received surprisingly little attention among marine environmental scientists. Before the EMERGE project there have been but a few peer-reviewed studies of scrubber water toxicity (Koski et al. 2017, Ytreberg et al. 2019). The ecotoxicological experiments carried out as part of EMERGE will hence contribute with much needed information on the environmental consequences of the introduction of scrubbers.

At present the debate outside the scientific community focuses mainly on the chemical content of the scrubber water and whether or not concentrations are below established threshold levels or water quality guidelines. Of central interest in these discussions is the IMO Guidelines for exhaust gas cleaning systems from 2015, which are the only discharge limits directed specifically at

scrubber water (Annex 1 Resolution MEPC.259(68) 2015)(IMO 2015). These guidelines include threshold levels only for pH, turbidity, nitrates, and what is referred to as total PAH, but un practice is what is called phenanthrene equivalents, a toxic equivalent factor focusing only on low molecular PAHs (Fisher et al. 2011). It is a well-known fact that the number of hazardous and potentially hazardous compounds deriving from combusted or non-combusted fuel that are present in scrubber water amounts to many hundreds or more. For most of these the toxicity is never tested although their chemical structures strongly suggest that toxic effects could be expected. Risk assessments of scrubber water solely based on chemical data hence run a serious risk of underestimating the true risk simply because most of the contaminants will not be accounted for. However, in an ecotoxicological test, with the whole scrubber effluent, the organisms will be exposed to all contaminants independently of whether or nor not we are aware of their presence in the scrubber water. Moreover, possible interactions between contaminants contributing to the overall ecotoxicity of the effluent can become visible in the overall adverse effect(s) on tested organisms.

For many compounds there are also national and international threshold levels that should not be exceeded in sea water to retain an ecological situation of no risk. The EU Water Framework Directive (2000/60/EC) has set this kind of threshold levels called Environmental Quality Standards (EQS) for in total 43 compounds (2013/39/EU) and some of them are present in wastewater from ship scrubbers. Another set of threshold levels for hazardous compounds in seawater are the Canadian Environmental Quality Guidelines (CEGQs). In many debates and reports on discharged scrubber water and the impact this may have on the environment, these threshold levels are referred to, and if they are not exceeded the discharges are considered relatively unproblematic (Clear Seas Centre for Responsible Marine Shipping 2022). However, in the EMERGE experiments the concentrations of individual PAHs and metals were many orders of magnitudes below these threshold levels yet, toxic effects were repeatedly detected.

An additional problem with the IMO guidelines for scrubber water is that they only consider the concentrations in the discharged water, not the total volumes being released. And these volumes have increased considerably since 2015 when the guidelines were introduced. Data based on ship activities in 2012 estimated the volumes of discharged scrubber water in the Baltic Sea to be around 1.5 million m^3 per year (Jalkanen et al. 2021), and with activity data updated to the situation 2018 the discharge volumes of scrubber water were estimated to be around 300 million m^3 per year (Ytreberg et al. 2022). Since 2018 the number of scrubbers has continued to increase, and the volumes of scrubber water discharge to the sea are hence massive. To not take this into account

when deciding on restrictions of the discharge of scrubber water to the sea must be consider as a serious shortcoming from all regulatory authorities involved.

5.3 Reflections on testing of scrubber water toxicity

5.3.1 Design of scrubber water experiments

When toxic effects of single compounds or of complex mixtures such as scrubber water are to be investigated, authorities frequently claim that the best way to do this is by running experiments according to standardised protocols (IMO 2019). Although this may be appropriate in some situations, e.g., for regularly recurring monitoring of a water body or a wastewater, it should not be confused with the far more extensive ambitions a research approach has for the same task. One of the key purposes of standardised protocols is that data should be completely comparable between tests. Therefore, requirements are very specific when it comes to physico-chemical conditions (temperature, salinity, pH etc.) and analysed end points. The recommended species are more often selected because they are easy to breed in the laboratory than because they are particularly sensitive to toxicants. Absolute comparability between tests is, however, not a key issue for ecotoxicological research, but the focus is rather on the researcher using his/her scientific skill and knowledge to explore what functions in an organism or in a community of organisms that might be sensitive to the compounds (or other stressors) of interest. Another important aspect is the ecological relevance of the choice of species and exposure conditions, which is often lost in standardised test. The relevance of the results may therefore be compromised. Still, there are a number of aspects that generally need to be "standardised" in order to maintain high-quality experimental research, and this includes issues like how the organisms should be treated to be in good condition, the expected fate of the tested compounds in the experimental set-up (hydrophobicity, vapour pressure, etc.) and statistical considerations. A good routine is hence to use a guide for quality criteria, such as the Criteria for Reporting and Evaluating Ecotoxicity Data (CRED), proposed by Moermond et al. (2016), as a checklist when designing an experiment.

In most ecotoxicological experiments with single compounds or mixtures, it is of utmost importance to handle the toxicants so that they as far as possible retain their original toxic potential throughout the experimental period. When working with scrubber water, there is a considerable risk of changing the chemical composition or characteristics of the effluent, and thus also the toxicity of the water, at any stage, from the sampling to the transport and storage of the samples, and even when the experiment is actually carried out. Exposure of scrubber water to air and/or UV

light will cause evaporation and photooxidation of many oil-derived compounds. If this exposure is combined with agitation of the water, the loss of many of the most toxic compounds may be considerable. This handling of scrubber water might partly explain the low toxicity of scrubber water observed in a study by DHI (2021). In this study scrubber water from several ships were mixed in a manner where it is likely that a significant portion of the more light-weight organic compounds have evaporated. Moreover, the study was conducted according to a standardised protocol that included a step of filtering the test water, and since many of the toxic compounds in scrubber water, both organic molecules and metals, have an affinity for particles this will most likely have led to a major reduction of the true toxicity of the water.

As clearly demonstrated by the highly significant data generated in EMERGE, it is very unfortunate that IMO recommends that only standardised protocols be used for ecotoxicological testing of scrubber water (IMO 2019). Considering the already existing quality assurance procedures for ecotoxicological data e.g., CRED analysis suggested by EC (2018) and the rigorous scientific peer-review process of scientific journals, the procedure recommended by IMO may actually produce data of lesser quality and relevance i.e., data unfit to protect marine life and ecosystems.

5.3.2 Comparison of toxicant concentration in scrubber water to background levels

Sea water sampled as a reference for background concentrations of contaminants must be collected away from areas with intense traffic, since the water in harbours and busy ship lanes for obvious reasons have permanently elevated pollutant concentrations. This was not always taken into account in the EMERGE project but is an important consideration for the future. Much of the data today on toxicant concentrations in seawater relative to those in scrubber water derives from samples collected onboard ships, before the water enters the scrubber (e.g., Magnusson et al. 2018). This water is contaminated both because it is collected in a ship lane, and because it has passed through pumps and tubes on the ship. This “inlet water” (water sampled onboard, prior to entering the scrubber) is for example found to be particularly high in copper compared to seawater outside ship lanes. Copper is used as anti-fouling paint on most ships today, and although there is little data available, it is most likely that the copper concentrations in ship lanes or harbours are elevated compared to the sea water elsewhere. Copper is also generally part of alloys used in piping. There are hence at least two important sources for the elevated copper concentrations in the inlet water used for scrubbing of exhaust gases. Even concentrations of oil-derived contaminants will most

likely be elevated in ship lanes and harbours, not least because this is where the scrubber water is discharged. The take home message is hence that much of the data on pollutant concentrations in "sea water" used for comparison with concentrations in scrubber water are not representative for sea water concentrations outside ship lanes and harbours.

6 CONCLUSIONS

Toxic effects of scrubber water on marine organisms were detected at much lower concentrations than in previous studies. Most sensitive among the tested endpoints were fertilisation and larval development of invertebrates. None of these parameters have previously been included in studies on scrubber water toxicity, but they are of central importance for the survival and existence of marine invertebrates and thus for maintaining sustainable marine food webs. Growth rate of microalgae and mortality of invertebrate species are toxicological endpoints commonly tested in standardised ecotoxicological tests; however, these parameters were found to be less sensitive to scrubber water.

All scrubber waters that were tested in the EMERGE experiments were treated with utmost care, from the sampling, to transport, storage and during the actual execution of the experiments, which is a guarantee for the validity of the obtained data. All physical disturbances of the water, like agitation and aeration will reduce the toxicity by causing evaporation of many of the more low-molecular oil-related compounds. Some protocols for standardised ecotoxicological tests states that the test water should be filtered before conducting the experiments, which, since a large part of the toxic compounds in scrubber water have a large affinity for particles, will inevitably lead to a reduced toxicity and an underestimation of the risk scrubber water pose on the marine environment. Another recommendation of standardised protocols is to adjust the pH so that the more acidic scrubber water has a higher pH resembling that of seawater. This may, however, affect the speciation and toxicity of metals and the effect of pH itself will be lost, thus changing the real toxicity of the scrubber water.

The overall outcome from the ecotoxicological activities of the EMERGE project is that there is an apparent risk that scrubber water may have a serious impact on the populations of key species of marine food webs, and that this should be taken into account in the continued debate about the future use of scrubbers.

Furthermore, we stress the unfortunate circumstance that IMO recommends using standardised protocols only for ecotoxicological testing of scrubber water (IMO 2019). Considering the already existing quality assurance procedures for ecotoxicological data and risk assessment e.g., CRED analysis suggested by EC (2018) and the rigorous scientific peer-review process of scientific journals, the procedure recommended by IMO may actually produce data of lesser quality and relevance, i.e., data unfit to protect marine life and ecosystems.

REFERENCES

Backhaus, T., J. Sumpter, and H. Blanck. 2008. On the Ecotoxicology of Pharmaceutical Mixtures. Pages 257-276 *in* K. Kümmerer, editor. Pharmaceuticals in the Environment: Sources, Fate, Effects and Risks. Springer Berlin Heidelberg, Berlin, Heidelberg.

Baden, S., B. Hernroth, and O. Lindahl. 2021. Declining Populations of Mytilus spp. in North Atlantic Coastal Waters—A Swedish Perspective. Journal of Shellfish Research **40**:269-296.

Beaugrand, G., K. M. Brander, J. A. Lindley, S. Souissi, and P. C. Reid. 2003. Plankton effect on cod recruitment in the North Sea. Nature **426**:661-664.

Beaugrand, G., and R. R. Kirby. 2010a. Climate, plankton and cod. Global Change Biology **16**:1268-1280.

Beaugrand, G., and R. R. Kirby. 2010b. Spatial changes in the sensitivity of Atlantic cod to climate-driven effects in the plankton. Climate Research **41**:15-19.

Chan, K. Y. K., M. A. Sewell, and M. Byrne. 2018. Revisiting the larval dispersal black box in the Anthropocene. Ices Journal of Marine Science **75**:1841-1848.

Clarke, K. R., and R. N. Gorley. 2006. PRIMER v6: UserManual/Tutorial. PRIMER-E, Plymouth.

Clear Seas Centre for Responsible Marine Shipping. 2022. Assessing Pollutants in Scrubber Discharge Water. Informing the Regulation of Ships' Exhaust Gas Cleaning Systems.

DHI. 2021. Ecotoxicity testing and risk assessment of wash water from open loop scrubbers. Prepared for the Exhaust Gas Cleaning Systems Association (EGCSA). Final report 11826102.

DNV-GL. 2022. Alternative Fuels Insight Platform (AFI) [Dataset]. Retrieved September 23, 2022, from https://store.veracity.com/da10a663-a409-4764-be66-e7a55401275a.

EC. 2018. Technical guidance for deriving environmental quality standards. Guidance Document no 27. Updated version 2018

Fisher, T. T., R. J. Law, H. S. Rumney, M. F. Kirby, and C. Kelly. 2011. Towards a scheme of toxic equivalency factors (TEFs) for the acute toxicity of PAHs in sediment. Ecotoxicology and Environmental Safety **74**:2245-2251.

Folt, C. L., and C. W. Burns. 1999. Biological drivers of zooplankton patchiness. Trends in Ecology & Evolution **14**:300-305.

Hassellöv, I.-M., D. R. Turner, A. Lauer, and J. J. Corbett. 2013. Shipping contributes to ocean acidification. Geophysical Research Letters **40**:2731-2736.

Heath, M. R., and R. G. Lough. 2007. A synthesis of large-scale patterns in the planktonic prey of larval and juvenile cod (Gadus morhua). Fisheries Oceanography **16**:169-185.

IMO. 2015. Annex 1 Resolution MEPC.259(68) 2015 Guidelines for Exhaust Gas Cleaning Systems. Marine Environment Protection Committee, International Maritime Organization.

IMO. 2019. Evaluation and harmonization of rules and guidance on the discharge of liquid effluents from EGCS into waters, including conditions and areas P. I. Sub-committee on pollution prevention and response.

Jalkanen, J. P., L. Johansson, M. Wilewska-Bien, L. Granhag, E. Ytreberg, K. M. Eriksson, D. Yngsell, I. M. Hassellöv, K. Magnusson, U. Raudsepp, I. Maljutenko, H. Winnes, and J. Moldanova. 2021. Modelling of discharges from Baltic Sea shipping. Ocean Sci. **17**:699-728.

Kase, R., M. Korkaric, I. Werner, and M. Ågerstrand. 2016. Criteria for Reporting and Evaluating ecotoxicity Data (CRED): comparison and perception of the Klimisch and CRED methods for evaluating reliability and relevance of ecotoxicity studies. Environmental Sciences Europe **28**:7.

Klimisch, H. J., M. Andreae, and U. Tillmann. 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul Toxicol Pharmacol **25**:1-5.

Koski, M., C. Stedmon, and S. Trapp. 2017. Ecological effects of scrubber water discharge on coastal plankton: Potential synergistic effects of contaminants reduce survival and feeding of the copepod *Acartia tonsa*. Marine Environmental Research **129**:374-385.

Lunde Hermansson, A., I.-M. Hassellöv, J. Moldanová, and E. Ytreberg. 2021. Comparing emissions of polyaromatic hydrocarbons and metals from marine fuels and scrubbers. Transportation Research Part D: Transport and Environment **97**:102912.

Magnusson, K., P. Thor, and M. Granberg. 2018. Scrubbers: Closing the loop Activity 3: Task 2 Risk Assessment of marine exhaust gas scrubber water. IVL Swedish Environmental Research Institute, Stockholm

MEPC-IMO. 2018. MEPC 74/INF.24 Annex: Report by the expert board for the environmental impact assessment of discharge water from Scrubbers (Japan).

Moermond, C. T. A., R. Kase, M. Korkaric, and M. Ågerstrand. 2016. CRED: Criteria for reporting and evaluating ecotoxicity data. Environmental toxicology and chemistry **35**:1297-1309.

Moldanova, J., E. Fridell, O. Popovicheva, B. Demirdjian, V. Tishkova, A. Faccinetto, and C. Focsa. 2009. Characterisation of particulate matter and gaseous emissions from a large ship diesel engine. Atmospheric Environment **43**:2632–2641.

Pineda, J., J. A. Hare, and S. Sponaugle. 2007. Larval Transport and Dispersal in the Coastal Ocean and Consequences for Population Connectivity. Oceanography **20**:22-39.

Seddiek, I. S., and M. M. Elgohary. 2014. Eco-friendly selection of ship emissions reduction strategies with emphasis on SOx and NOx emissions. International Journal of Naval Architecture and Ocean Engineering **6**:737-748.

Stige, L. C., G. Ottersen, D. O. Hjermann, P. Dalpadado, L. K. Jensen, and N. C. Stenseth. 2011. Environmental toxicology: Population modeling of cod larvae shows high sensitivity to loss of zooplankton prey. Marine Pollution Bulletin **62**:395-398.

Thor, P., and S. Dupont. 2018. Ocean acidification.*in* S. M and M. T, editors. Handbook on marine environment protection. Springer, Berlin.

Thor, P., M. E. Granberg, H. Winnes, and K. Magnusson. 2021. Severe Toxic Effects on Pelagic Copepods from Maritime Exhaust Gas Scrubber Effluents. Environmental Science & Technology **55**:5826-5835.

Ytreberg, E., K. Hansson, A. L. Hermansson, R. Parsmo, M. Lagerström, J.-P. Jalkanen, and I.-M. Hassellöv. 2022. Metal and PAH loads from ships and boats, relative other sources, in the Baltic Sea. Marine pollution bulletin **182**:113904.

Ytreberg, E., I.-M. Hassellöv, A. T. Nylund, M. Hedblom, A. Y. Al-Handal, and A. Wulff. 2019. Effects of scrubber washwater discharge on microplankton in the Baltic Sea. Marine Pollution Bulletin **145**:316-324.

Ytreberg, E., A. Lunde Hermansson, and I. M. Hassellöv. 2021. Database and analysis on waste stream pollutant concentrations, and emission factors; EMERGE report Deliverable 2.1.

Zhao, J., Y. Zhang, J. Chang, S. Peng, N. Hong, J. Hu, J. Lv, T. Wang, and H. Mao. 2020. Emission characteristics and temporal variation of PAHs and their derivatives from an ocean-going cargo vessel. Chemosphere **249**:126194.

APPENDICES

Appendix 1: Ecotoxicological results from UV

Marco Picone

In the present report we describe the toxicity testing performed with scrubber watersamples at the Venice University for the North Adriatic Sea (NAS) case study.

Scrubber water used

Toxicity tests were performed on two different scrubber water samples. A first sample was obtained from the Chalmers University of Technology and seawater collected in the Gulf of Venice was used to produce the scrubber washwater. A second scrubber water sample was obtained from the LEO C campaign. The sample was collected in the Mediterranean Sea (IDpoint 10B_SCRW). The scrubber water from LEO C collected for the NAS case study was relatively enriched (i.e., concentration > 75th percentile as compared with the other samples collected during the LEO C campaign) in alkylated phenanthrenes (phenanthrene-1C, phenanthrene-2C, phenanthrene-4C), and some trace elements, including V, Mn, Ni, Cd, and Pb). Artificial scrubber water obtained from Chalmers was characterized by lower concentrations of V and PAHs compared to LEO C scrubber water, but relevantly higher concentration of Cr (approx. 8-fold higher), Fe (approx. 6-fold higher), Cu (approx. 10-fold higher).

Methods

The scrubber water obtained from Chalmers was tested using a suite of four toxicity testing,using acute, early-life stages, and chronic exposure:

1) The Microtox test with *Aliivibrio fischeri*. The test aims at verifying inhibitory effects onthe bacterial bioluminescence after 5, 15 and 30 minutes of exposure, according to the ISO 11348 standard method (ISO, 2007). The test was performed at 15°C on the following scrubber water concentrations (%): 0.01, 0.1, 1, 2, 5, 10, 20, 40.
2) The 48-h acute lethality test with the copepod *Acartia tonsa*. The test was performed according to the ISO 14669 standard guide (ISO, 1999). The test was performed at 20°Cusing a 20‰ salinity medium as dilution water. The following scrubber water concentrations were tested (%): 0.01, 0.1, 1, 2, 5, 10, 20, 40.
3) The larval development test with the copepod *A. tonsa.* The test aims at identifying

detrimental effects on hatching, early-life stage survival and larval development (to thecopepodite-I stage) after 5 days of exposure to the scrubber water. The test was performed according to the ISO 16778 standard guide (ISO, 2015a), modified according to Picone et al. (2022, 2021). The test was performed at 20°C using a 20‰ salinity medium as dilution water. The following scrubber water concentrations were tested (%): 0.01, 0.1, 1, 2, 5, 10, 20, 40.

4) The long-term exposure of *A. tonsa*. The test was performed according to Picone et al. (2022 submitted). The test aimed to identify effects on the parental generation (F_0), using egg production as the endpoint (after 13 days of exposure), and on the offspring (generation F_1), using hatching ratio, early life stage mortality, and larval development ratio as endpoints after a larval development test performed as described at point 3), after 21 days of exposure for F_0 and 5 days of exposure for F_1. The test was performed at 20°C using a 20‰ salinity medium as dilution water. The 0.01%, 0.1%, and 1% scrubber water concentrations were tested.

The LEO C scrubber water was tested using an enlarged suite of bioassays, using acute, early-life stages, and chronic exposure, including:

1) The Microtox test, according to the above-described conditions.

2) The 96-h algal growth test using the diatom *Phaeodactylum tricornutum* and the chlorophyte *Dunaliella tertiolecta.* The test aims at identifying detrimental effects on thealgal growth rate. The tests were performed according to the ASTM E1218 standard guideline (ASTM, 2021) at 20°C, using Guillard's F/2 medium (Guillard and Ryther, 1962) as dilution water (salinity 34‰). The scrubber water concentration tested were the following: 5%, 10%, 20%, and 40%.

3) The larval development test with *M. galloprovincialis.* The test was performed according to an internal protocol based on the standard ISO 17244 (ISO, 2015b). The test aimed to verify the effects on larval development by discriminating the normally developed prodissoconch-I larvae (D-shaped larvae) from abnormal prodissoconch-Ilarvae and delayed larval stages (trochophore larvae, gastrulae). The test was performed at 18°C for 48 hours. The tested concentrations (%) were the following: 0.001, 0.01, 0.1, 1, 2, 5, 10, 20, 40.

4) The 48-h acute lethality test with the copepod *Acartia tonsa*, according to the above-described conditions.

5) The larval development test with the copepod *A. tonsa*, according to the above-described conditions, with the addition of the 0.001% test concentration.

6) The long-term exposure of *A. tonsa*, according to the above-described conditions, withthe addition of the 0.001% test concentration.

Physico-chemical parameters measured for each test and treatment are reported in the appendix. Scrubber water samples were not buffered before testing.

Results

Chalmers scrubber water

The scrubber water obtained from Chalmers was acutely toxic to bacteria and copepods only at the highest tested concentrations (NOEC = 10%, LOEC = 20% for the Microtox test; NOEC = 20%, LOEC = 40% for the acute test with *A. tonsa*). The EC_{10} were 24.8% and 36.4% for *A. fischeri* and *A. tonsa*, respectively. The effects on early-life stages of *A. tonsa* were muchmore pronounced, but only for the larval development endpoint, for which significant effects were observed at all the tested concentrations (NOEC < 0.01%; LOEC = 0.01%).

Effects on hatching (NOEC = 20%, LOEC = 40%) and larval survival (NOEC = 10%, LOEC =20%, EC10 = 11.8%) were similar to that observed for adult copepods and bacteria. An $EC_{50}$was calculable only for early-life stage survival (EC_{50} = 19.6%) Only the egg production endpoint was measured in the long-term test with copepods. The data evidenced a strong effect on the reproduction ability of *A. tonsa*, with significant inhibition observed at the lowest tested concentration (NOEC < 0.01%, LOEC = 0.01%).

LEO C scrubber water

The scrubber water significantly reduced the bioluminescence compared to the negative control only at high concentrations (NOEC = 10%, LOEC = 20%, EC_{10} = 23% at 30 minutes). Similarly, also the inhibition of algal growth occurred at relatively high scrubber water concentrations, with effect more marked on *D. tertiolecta* (NOEC = 10%, LOEC = 20%, EC_{10} =15%) than on *P. tricornutum* (NOEC = 20%, LOEC = 40%, EC_{10} = 34%). Lethal effects of scrubber water on *A. tonsa* adults showed differences between specimens cultured at 20‰ and 30‰ salinity with NOECs and LOECs evidencing specimens cultured at 20‰ salinity (NOEC = 5%, LOEC = 10%, EC_{10} = 8%, EC_{50} = 11%) being slightly more sensitive to scrubber water than individuals grown at 30‰ salinity (NOEC = 10%, LOEC = 20%, EC_{10} = 10%, EC_{50}= 17%).

The larval development of *M. galloprovincialis* was severely affected by exposure to the scrubber water. All larvae were at the early trochophore stage at 10% scrubber water concentration. At 20% scrubber water, the development was blocked at the 4-cell stage,

while at 40% scrubber water, only the first polar body was observed. The EC_{50} was calculated at 6% scrubber water, while NOEC, LOEC and EC_{10} were estimated at 0.1%, 1% and 4.9%, respectively. Similarly, the larval development of copepods was severely inhibited at a concentration as low as 2% scrubber water; the EC_{50} was estimated at 1.5% scrubber water, while NOEC, LOEC and EC_{10} were estimated at 1%, 2% and 1,1%, respectively. On the contrary, hatching and larval survival showed a lesser sensitivity to thetoxic action of the scrubber water. They provided results similar to the acute tests with bacteria and copepods (NOEC = 10%, LOEC = 20%, EC_{10} = 9% for both hatching and survival).

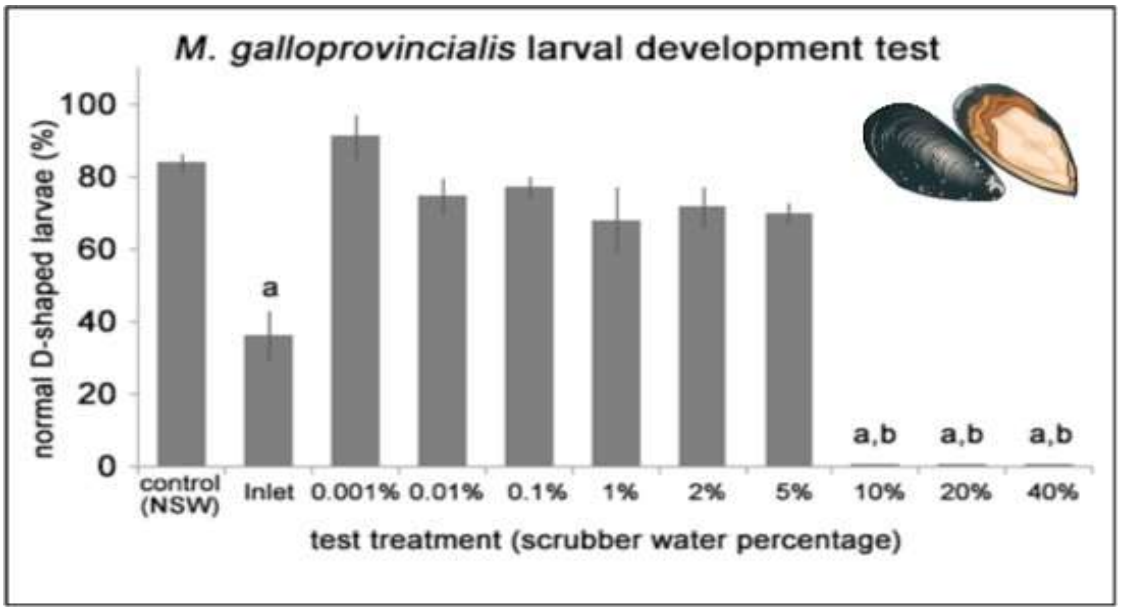


Figure A-1. Results of the early-life stage test with *M. galloprovincialis* performed on scrubber water, inlet water, and North Adriatic Sea water (NSW). a = treatments with response significantly lower than the negative control (NSW); b = treatments with response significantly lower than negative control and inlet water.

As concern the long-term exposure, egg production by F_0 generation showed a decreasing trend up to a scrubber water concentration of 0.1%, and then it increased at 0.1% and 1% concentrations. However, egg production was lower in all treatments than in the negative control (NOEC = 0.001%, LOEC = 0.01%). The hatching ratio and early-life stage survival of generation F_1 were unaffected by the exposure to scrubber water. Conversely, the larval development of the F_1 generation showed a U-shaped trend similar to that observed for eggproduction by the parent generation. LDR decreased almost linearly from the negative control up to 0.1% scrubber water concentration, and then it increased at 1% scrubber water(NOEC = 0.01%, LOEC = 0.1%).

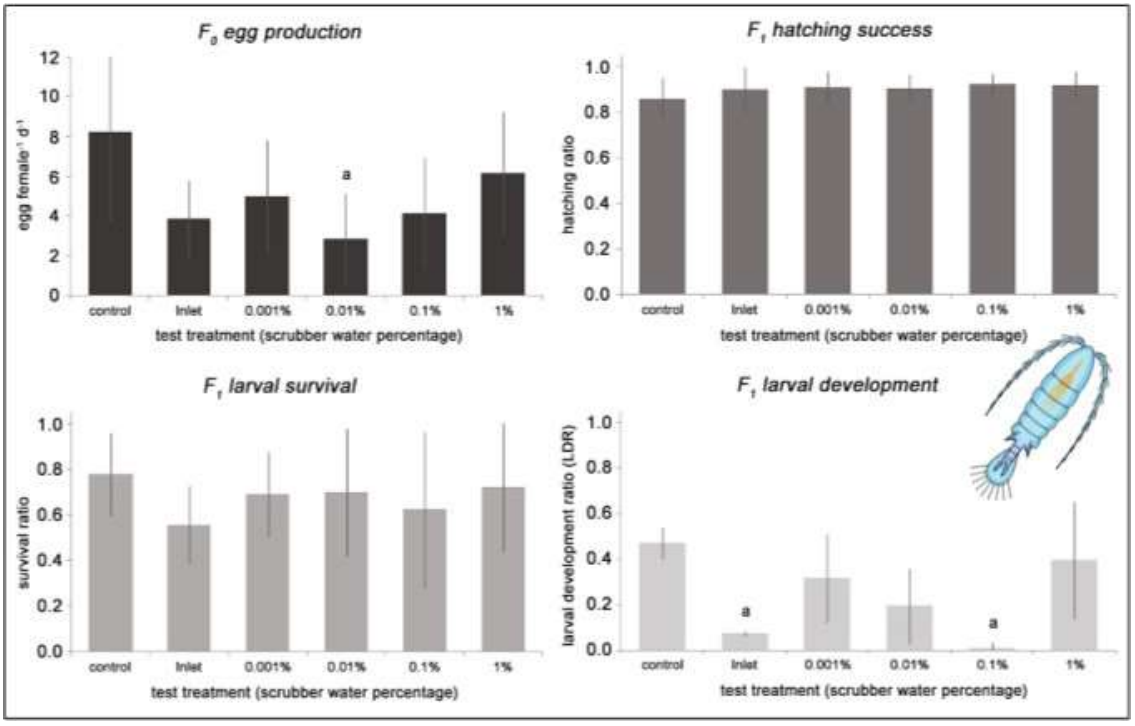


Figure A-2. Results of the long-term exposure test with *A. tonsa*: egg production by F_0 generation (panel A), hatching of F_1 generation (panel B), larval survival for F_1 (panel C), and larval-development of F_1 generation performed on scrubber water and inlet water. Lower case letters indicate significant differences: a = treatments with response significantly lower than the negative control. Error bars designate standard deviations.

Discussion and conclusion

The exposure of planktonic indicators to scrubber water evidenced concentration-dependent effects on most of the explored endpoints. Acute effects on bacteria, algae, and copepods were observed at relatively high concentrations. Conversely, effects on larval stages of mussel and copepods occurred at scrubber water concentrations sensibly lower than acute effects, with larval development of *A. tonsa* being the more sensitive endpoint (EC_{10} = 1.1%). Similarly, also the onset of effects on mussel development occurred at a scrubber water concentration lower than the acute effects (EC_{10} = 4.9%). The long-term exposure of *A. tonsa* produced the lowest effect concentrations since both egg production bygeneration F_0 and the larval development of generation F_1 were significantly affected at a scrubber water concentration of < 0.1%.

Since single contaminants were generally below the adverse effect levels for the planktonic bioindicators used in the study, the toxicity of scrubber water was most probably due to synergistic effects of the chemical mixture and physicochemical properties (i.e., pH) of the scrubber water. Nevertheless, for many contaminants of potential concern, such as V and alkylated PAHs, there is the need to increase the dataset of adverse effects levels since for many endpoints (i.e. larval development with both bivalves and copepods and effects on

reproduction of copepods) such information is not available in the literature.

References

ASTM, 2021. Standard guide for conducting static toxicity tests with microalgae - ASTM E1218. https://doi.org/10.1520/E1218-21

Guillard, R.R., Ryther, J.H., 1962. Studies of marine planktonic diatoms: I. *Cyclotella nana*Hustedt, and *Denotula confervacea* (Cleve) GRAN. Can. J. Microbiol. 8, 229–239. https://doi.org/10.1139/M62-029

ISO, 2015a. Water quality — Calanoid copepod early-life stage test with *Acartia tonsa* - ISO16778:2015.

ISO, 2015b. Water quality — Determination of the toxicity of water samples on the embryo- larval development of Japanese oyster (*Crassostrea gigas*) and mussel (*Mytilus edulis* or *Mytilus galloprovincialis*). ISO 17244:2015.

ISO, 2007. Water quality — Determination of the inhibitory effect of water samples on thelight emission of *Vibrio fischeri* (Luminescent bacteria test) — Part 3: Method using freeze-dried bacteria - ISO 11348-3:2007.

ISO, 1999. Water quality — Determination of acute lethal toxicity to marine copepods (Copepoda, Crustacea) ISO 14669:1999.

Picone, M., Distefano, G.G., Marchetto, D., Russo, M., Baccichet, M., Zangrando, R., Gambaro, A., Volpi Ghirardini, A.V., 2022. Inhibition of larval development of marine copepods *Acartia tonsa* by neonocotinoids. Toxics 10, 158. https://doi.org/10.3390/TOXICS10040158

Picone, M., Distefano, G.G., Marchetto, D., Russo, M., Vecchiato, M., Gambaro, A., Barbante, C., Volpi Ghirardini, A., 2021. Fragrance materials (FMs) affect the larval development of the copepod *Acartia tonsa*: An emerging issue for marine ecosystems. Ecotoxicol.

Environ. Saf. 215, 112146. https://doi.org/10.1016/j.ecoenv.2021.112146

Picone, M., Distefano, G.G., Marchetto, D., Russo, M., Baccichet, M., Brusò, L., Zangrando, R., Gambaro, A., Volpi Ghirardini, A.V., (submitted, minor revision due). Long-term effects of neonicotinoids on reproduction and offspring development in the copepod *Acartia tonsa.* Marine Environmental Research.

Appendix 2: Ecotoxicological results from UoS

Zapata-Restrepo L.M., Hudson M.D., and Williams I.D

In the present report we describe the ecotoxicological tests of open-loop scrubber water from exhaust gas cleaning systems from ships on blue mussels (*Mytilus edulis*) and sea-urchin (*Psammechinus miliaris).* For this, chronic toxicity tests were performed using a fertilization test anda larval development bioassay at the University of Southampton (UoS) for the Solent case study.

Scrubber water used

Two different scrubber water samples were used in this study:

1) The Chalmers University of Technology produced scrubber washwater using seawater collected in the Gulf of Venice.
2) The scrubber effluent collected from an open-loop system on board the DANAOS CatherineC container ship during an on-board campaign at the English Channel at Sea (Sampling site:1, location: 51◦04.6'N; 001◦37.9'E; sample name: 1B_SCRW) on the 18th of August 2021.

PAHs and heavy metals were measured in scrubber waters prior to the experiments. PAHs analyses were carried out in filtered and unfiltered water samples by Catalan Institute for Water Research (ICRA). Scrubber waters were filtered, and heavy metal analysis were carried out in both solution andfilters at the National Oceanography Centre Southampton (NOCS).

The scrubber water from DANAOS Catherine C collected for the Solent case study presented higherconcentrations for all the 16 PAHs and 13 AlkylPAHs analysed. According to these results, PAHs and heavy metals are lost after the filtration process because they remain adsorbed to the particulate matters present in the samples. For this reason, the scrubber water used for experimental treatments was unfiltered in all cases.

Methods

Scrubber water effluents obtained from Chalmers and DANAOS were tested using and two chronic toxicity tests (fertilization success and larval development) on blue mussels (*Mytilus edulis*) and sea-urchin (*Psammechinus miliaris).* DANAOS scrubber water samples did not require any pH correction.Chalmers scrubber water was tested with and without pH correction.

Fertilization assay on *M. edulis*:

Animals were induced to spawn by thermal shock and the best three females and males were chosen according to the eggs and sperm quality. The sperm was exposed during 30min to the different scrubber water dilutions. Then the oocytes were added and exposed during 30min at roomtemperature. Formalin was added to block embryonic development at an early stage. At least 100 embryos from each replicate were scored for percentage fertilization based on the number of cells showing cleavage. The scrubber water dilutions tested 0.001, 0.01, 0.1, 1, 2, 5, 10, 20, 40 and 100% of the original sample. Five replicates were used for each dilution.

Larval development test on M. edulis:

An internal protocol based on the standard ISO 17244 (ISO, 2015b) was followed. The scrubber water dilutions tested 0.001, 0.01, 0.1, 1, 2, 5, 10, 20, 40 and 100% of the original sample. Briefly, after checking the good quality of fresh eggs and sperm, and not later than 60 min of spawning, active sperm from three males is added to the solution of fresh sea water containing eggs from threefemale at ambient temperature and with continuous aeration. Fertilization was accomplished at a ratio of between 100 and 200 sperm egg-1 in a measuring cylinder. Once polar bodies were detectedat 20-30 min after fertilization, the activated egg suspensions at a density of 20–200 eggs cm-2 were distributed into flat-bottomed glass dishes or trays containing the test solutions. The fertilized eggs were kept undisturbed in the dark without aeration and food addition for up to 72h at 16°C. After the incubation period a few drops of 10% buffered formalin were added to each vessel to fix and preserve the larvae. Random samples of 100 larvae per replicate were counted distinguishing between normal larvae (D-shaped) and abnormalities (malformed larvae and pre-larval stages).

Larvae were considered abnormal if they present at least shell abnormality, mantle abnormality, segmentation abnormalities, and/or empty shell. The acceptability of test results was based on negative control for a percentage of normal D-shaped larvae ≥80%. EC_{50} is calculated on the basis ofthe percentage of normally developed D-stage larvae. Copper sulfate pentahydrate ($CuSO_4 \cdot 5H_2O$) is the recommended reference substance. The test concentrations were included in the range 0 µg/L to 100 µg/L of $CuSO_4 \cdot 5H2O$. Five replicates were used for each dilution. The test was performed in triplicate, in three successive series, using three batches of fertilized eggs.

Fertilization assay on *P. miliaris*:

The test was performed according to Environment Canadá (2014). Animals were induced to spawn by thermal shock and the best three females and males were chosen according to the eggs and sperm quality. The sperm was exposed during 30min to the different scrubber water

dilutions. Then the oocytes were added and exposed during 40min at room temperature. Formalin was added to block embryonic development at an early stage. At least 100 embryos from each replicate were scored for percentage fertilization based on the number of cells showing cleavage. The scrubber water dilutions tested 0.001, 0.01, 0.1, 1, 2, 5, 10, 20, 40 and 100% of the original sample. Five replicates were used for each dilution.

Larval development test on *P. miliaris*:

The larval development test on *P. miliaris* was performed with some minor adaptations of the protocol described for *M. edulis*. Briefly, after checking the good quality of fresh eggs and sperm, and not later than 2h of spawning, active sperm from three males is added to the solution of fresh sea water containing eggs from three female at ambient temperature and with continuous aeration.Fertilization was accomplished at a ratio of between 100 and 200 sperm egg-1 in a measuring cylinder. Once the fertilization ring was detected at 20-30 min after adding sperm and eggs, the activated egg suspensions at a density of 20–200 eggs cm-2 were distributed into flat-bottomed glass dishes or trays containing the test solutions. The fertilized eggs were kept undisturbed in the dark without aeration and food addition for up to 72h at 15°C. After the incubation period a few drops of 10% buffered formalin were added to each vessel to fix and preserve the larvae. Random samples of 100 larvae per replicate were counted distinguishing between Pluteus larvae considered as normal growth or any trochophore larvae, gastrulae, blastulae, morula, or split egg considered aslarvae with abnormal. The acceptability of test results was based on negative control for a percentage of normal larvae ≥80%. EC^{50} is calculated on the basis of the percentage of normally developed larvae. Copper sulfate pentahydrate ($CuSO_4 \cdot 5H2O$) is the recommended reference substance. The test concentrations were included in the range 0 µg/L to 100 µg/L of $CuSO_4 \cdot 5H_2O$. Five replicates were used for each dilution. The test was performed in triplicate, in three successiveseries, using three batches of fertilized eggs.

Results

Effects of scrubber water on fertilization success

The scrubber water obtained from Chalmers and DANAOS produced sublethal negative effects on fertilization success in both species (Figure A-3). In this test, the sea urchin embryos were more sensitive (NOEC = 0.1%; LOEC = 1%; EC_{10}=0.12%; EC_{50}=4.31%) than the blue mussel embryos (NOEC =1%; LOEC = 2; EC_{10}=0.49%; EC_{50}=156%) when exposed to DANAOS scrubber water. The same behaviour was observed when the test was performed with Chalmers scrubber water with or without pH correction. The sea urchin embryos were more sensitive (NOEC = 0.001%; LOEC = 0.01%)than the blue mussel embryos (NOEC = 0.1%; LOEC =

1%) when exposed to Chalmers scrubber water. When the pH of Chalmers scrubber water was corrected the toxicity of the scrubber water decreased for both species (*M. edulis*: Without pH correction: EC_{50}=40.89%; with pH correction: EC_{50}=85.32%; *P. miliaris*: Without pH correction: EC_{50}=1.2%; with pH correction: EC_{50}=1.49%).

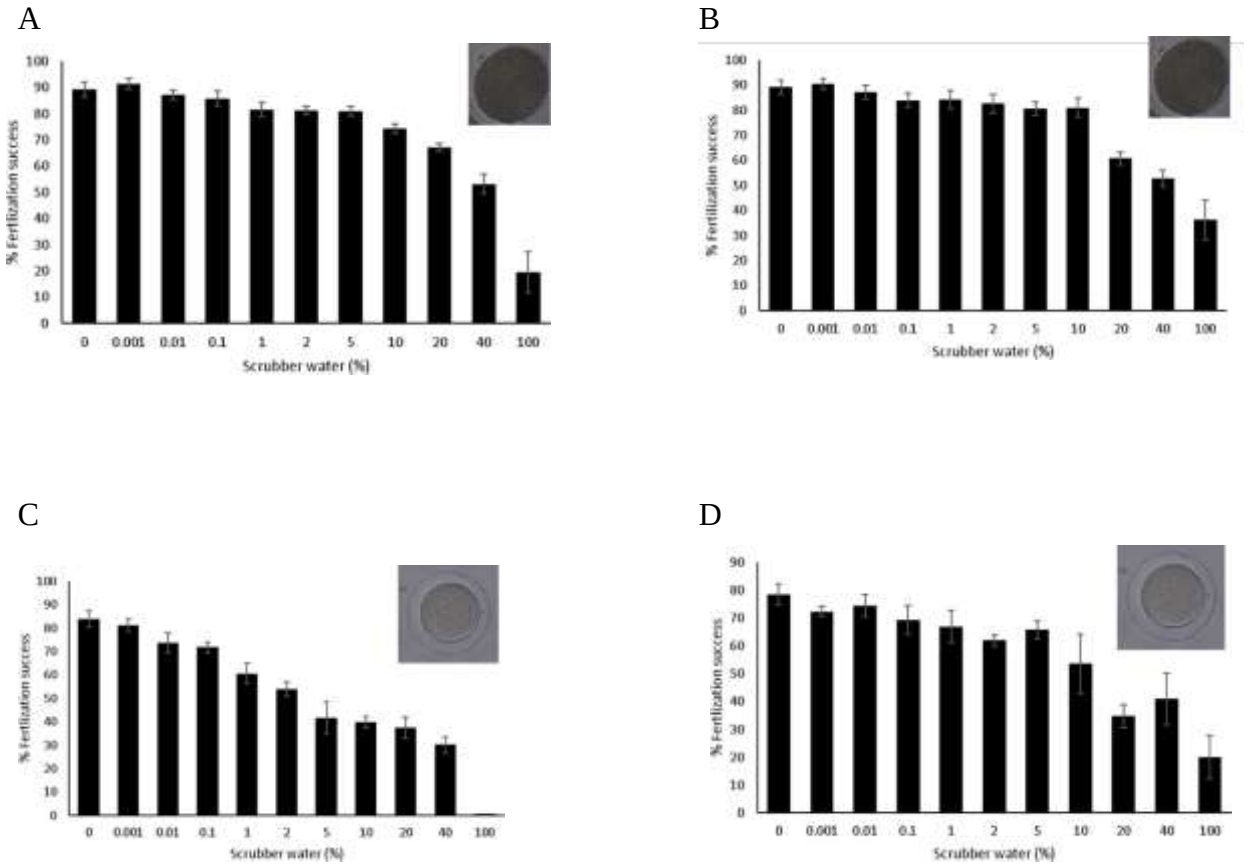


Figure A-3. Percentage of fertilization success in two species after exposure to scrubber water. *M. edulis* exposed to (A)Chalmers scrubber water and (B)Danaos scrubber water; *P. miliaris* exposed to(C)Chalmers scrubber water and (D)Danaos scrubber water.

Effects of scrubber water on larval development

The effects on early-life stages showed that larval of both species are sensitive to low concentrationsof scrubber water (Figure A-3). At 72h of exposure some abnormalities observed in *P. miliaris* larvae included additional crossbarred body rod, missing or shorter arms, and apically crossed body rod (Figure A-5). In *M. edulis* hypertrophy of the mantle and hinge abnormality were the most common larvae abnormalities found at 72h of exposure to both scrubber waters. In this case, blue mussel larvae were much more sensitive (NOEC = <0.001%; LOEC = 0.001%; EC_{10}=0.27%; EC_{50}=9.27%) than sea urchin larvae (NOEC = 0.01%; LOEC = 0.1; EC_{10}=0.15%; EC_{50}=1.54%) when exposed to DANAOS scrubber water. The same behaviour was observed when the test was performed with Chalmers scrubber water with or without pH correction. The blue mussel larvae were more sensitive (NOEC = <0.001%; LOEC

= 0.001%) than the se urchin larvae (NOEC = 0.001%; LOEC = 0.01%) when exposed to Chalmers scrubber water without pH correction. When the pH of Chalmers scrubber water was corrected the toxicity of the scrubber water slightly decreased for both species.

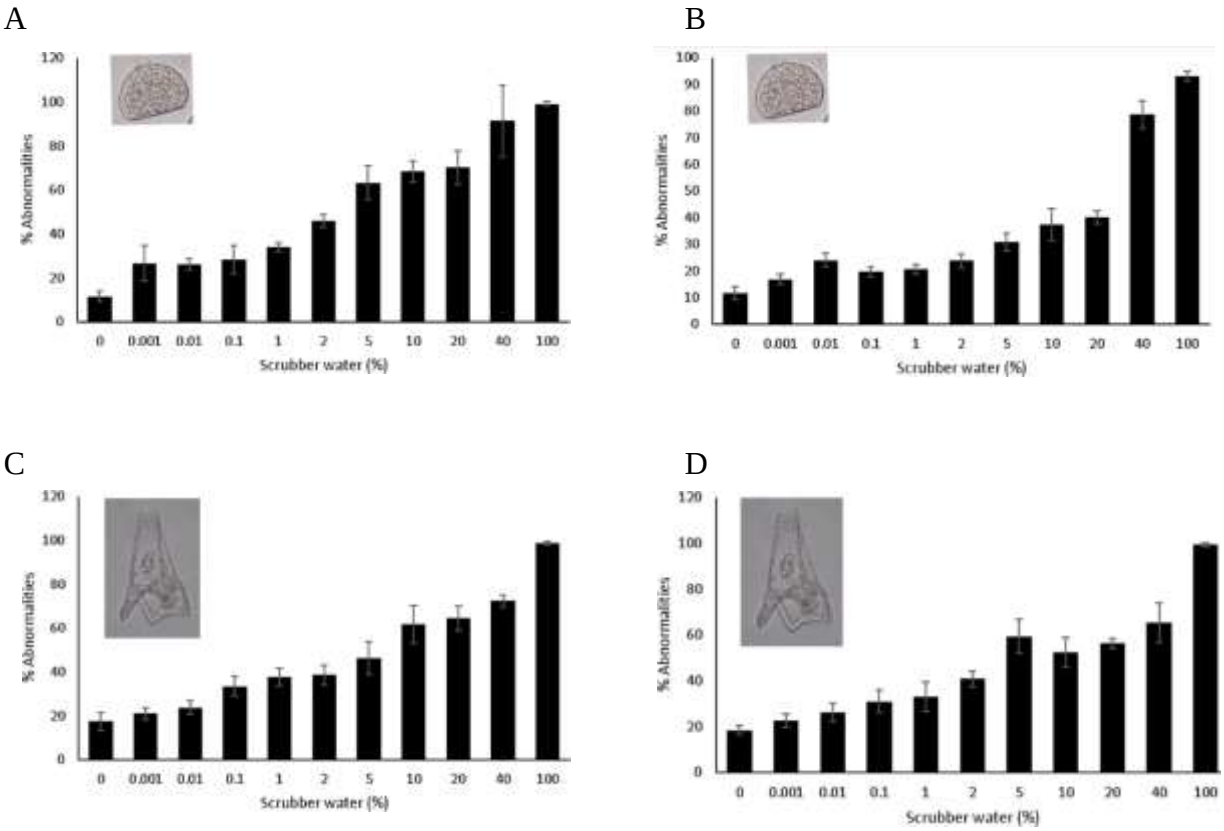


Figure A-4. Percentage of abnormal larvae in two invertebrate species after 72h exposure to scrubberwater. *M. edulis* exposed to (A)Chalmers scrubber water and (B)Danaos scrubber water; *P. miliaris* exposed to (C)Chalmers scrubber water and (D)Danaos scrubber water.

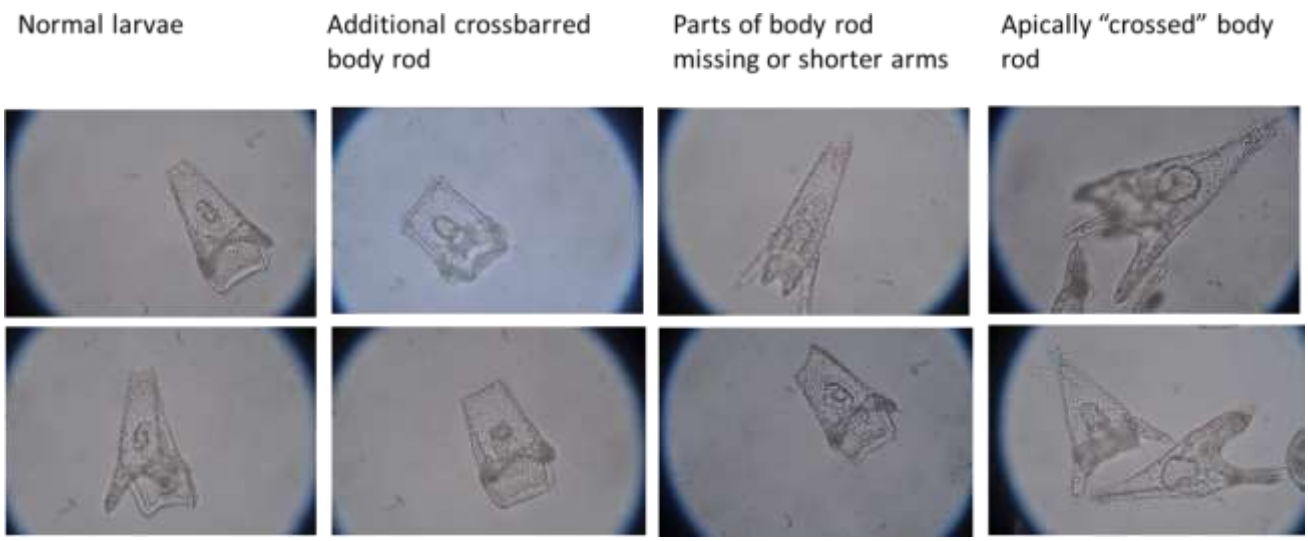


Figure A-5. Some abnormalities observed in *Psammechinus miliaris* larvae at 72h of exposure to scrubber water

Figure A-6. Some abnormalities observed in *Mytilus edulis* larvae at 72h of exposure to scrubber water.Arrows: hypertrophy of the mantle. Arrowhead: hinge abnormality.

Short discussion and conclusion

The exposure of *M. edulis* and *P. miliaris* embryos and larvae to different dilutions of scrubber watershowed a concentration-dependent effect. The adverse effects on fertilization success and larvae development were observed even at the lowest dilutions of scrubber water.

There is a differential sensitivity between species and life-stages to the exposure to scrubber water. Sea urchin embryos seemed to be more sensitive to this exposure than blue mussel embryos.

However, the pelagic larvae of *M. edulis* was more sensitive to the exposure to scrubber water.

To study the effect of scrubber water on different organisms it is necessary to consider some relevant aspects such as the pH correction and the filtration of samples. These processes can changethe behaviour of some of the compounds in the scrubber water. It was evident that the pH correction decreased the toxicity of the scrubber water. This was expected considering that a changein pH modifies the chemical form, solubility, and availability of some compounds present in the mixture. In the same manner, the filtration of scrubber water prior to the experiments can lead to the loss of some compounds that can remain adsorbed to the particulate matter present in the samples.

Despite the lowest concentration of most of the PAHs and heavy metals analysed in Chalmers scrubberwater it showed a higher potential to be toxic for both species, either for fertilization or the larval development. This confirms the potential of these effluents to be harmful to marine organisms and the need to study the interactions and agonisms and/or synergisms caused by compounds present incomplex mixtures such as scrubber water samples.

References

Environment Canada 2014. Reference Method for Measuring the Toxicity of Contaminated Sedimentto Embryos and Larvae of Echinoids (Sea Urchins or Sand Dollars) 1/RM/58 July

2014. Science and Technology Branch Environment Canada

ISO, 2015b. Water quality — Determination of the toxicity of water samples on the embryo-larval development of Japanese oyster (*Crassostrea gigas*) and mussel (*Mytilus edulis* or *Mytilus galloprovincialis*) ISO 17244:2015.

Appendix 3: Ecotoxicological results from IVL

Maria Granberg, Chiau Yu Chen & Kerstin Magnusson

Scrubber water used

The scrubber water used in the experiments was collected from the DANAOS ship Leo C at the northernmost station 1B (1B_SCRW) in the English Channel. This station was chosen due to its proximity to the Öresund case study area. Metal analyses were not available at thetime of writing since the original sample vials were broken upon arrival at ICRA. Metal analysis will hence be delayed. The PAH and alkylated PAH contents of the scrubber water from station 1B was very similar to most of the other scrubber waters. The pH of the scrubber water was 4.

Methods

After collection onboard ship the scrubber water was stored cool (+5 °C) and dark in acid (HCl) and acetone washed 5l glass flasks until arrival in port. The samples were transported to Kristineberg Marine Research Station, Sweden under cool (+5 °C) and dark conditions and kept under the same conditions at the research station until use. The period from collection on board until experimental start amounted to approximately 3 months. All material (aluminum foil, glassware, or plastic) coming into contact with scrubber water wasacid (HCl) and acetone washed or muffled (500 °C, 4 h) to remove traces of organic matter and contaminants prior to use.

Ecotoxicological experiments were conducted investigating the effects of scrubber water on various developmental stages of the green sea urchin (*Strongylocentrotus droebachiensis*), collected in Skagerrak on the Norwegian coast. Two types of experiments were carried out to investigate larval development and fertilization success. The chronic larval development exposures were carried out for 11 days at ambient temperature, +15 °C, under semi-static conditions in darkness, while the acute fertilization success exposures were carried out by exposing eggs and activated sperm to scrubber water for 15 minutes at ambient temperature, +15 °C, in darkness. Several endpoints were measured to quantify larval development and of the obtained results data on larval deformation (Figure A-9) was found to bemost suitable for calculating LOEC, NOEC and EC10 values. Fertilization success was determined by recording the presence or absence of the fertilization membrane (Figure A-10 and Figure A-11).

Spawning and fertilization

Adult sea urchins were kept in aquaria with running seawater (32 PSU) at ambient temperature (+15 °C) and fed intermittently for several weeks before experimental start. Spawning was induced in ripe sea urchins by injecting 0.5 M potassium chloride (KCl) solution (37 mg/ml in filtered seawater, FSW) into the coelomic cavity through the peristomal membrane. Female sea urchins were placed upside down onto 100 mL Erlenmeyer flask filled to the rim with filtered seawater (0.45 µm, 32 PSU) for eggs to aggregate at the bottom. Sperm were collected by pipetting freshly expelled semen into Eppendorf tube placed on ice. The egg suspension and semen were subsampled for density determination, compatibility test was done by mixing eggs with activated sperm and observed under microscope. If compatible, a larger batch of eggs were fertilized, and immediately used for the larval development tests.

Larval development experiment

Scrubber water was diluted with filtered seawater (FSW; 0.45 µm, 32 PSU) to achieve the following test concentrations 10%, 5%, 2%, 1%, 0.1%, 0.01%, 0.001%, 0.0001% and 0% (control). 305 – 315 (mean = 308) ml of experimental water was added to blue cap flasks andsubsequently a specific number of fertilized eggs was added to each flask. Larval development was monitored each day for 11 days from embryo to the pluteus larval stage. Larvae were fed microalgae (*Rhodomonas sp.*) every third day from day 6 and onward.

Exposure water was changed every fourth day by carefully wet-sieving out the larvae (20 µm pore size) and gently rinsing them back into their respective flasks after exposure

water replacement. Effects of scrubber water on *Rhodomonas sp.* as live feed was not tested since Koski et al. (2017) have shown limited and late onset of effects of scrubber water on the species. Each day a sample of 10 ml was retrieved from the experimental flasks and fixed with two drops of paraformaldehyde (4% in FSW). A minimum of 10 larvae per replicate was photographed under a microscope (Leica LEITZ DMRBE, 301-371.011) to determine body length, total length, body rod length, postoral rod length, posterolateral rod length and stomach dimensions using ImageJ (Figure A-7 right, explanations of the terms in the legend). Each measured larva was characterized either as normal or abnormal based on the general morphology (Figure A-8).

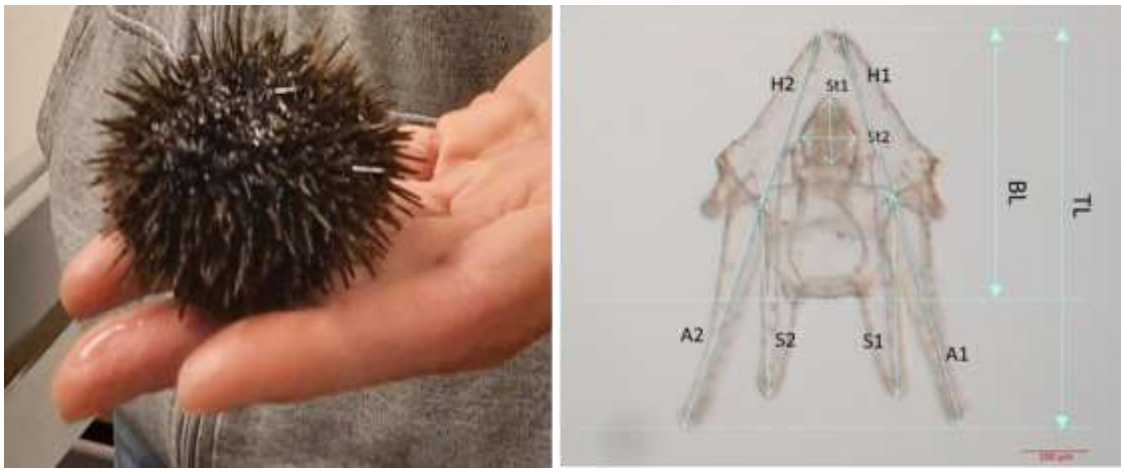


Figure A-7. The green sea urchin (*Strongylocentrotus droebachiensis*) as benthic adult (left) and pelagic larvae (right). Right picture shows morphometric measurement of sea urchin larvae. TL: total length; BL: body length; H1 and H2: left and right body rod length (BRL) respectively; A1 and A2: left and right posterolateral arm length (PLL) respectively; S1 and S2: left and right postoral rod length (POL) respectively; St1 and St2: Stomach length and width respectively. Image taken from larvae at Day 11 at 0.001% exposure.

Fertilization experiment

1.5 ml of scrubber water was added to 2 ml glass crystallization vials with the following experimental concentrations: 50%, 20%, 10%, 5%, 2%, 1%, 0.1%, 0.01%, 0.001% and 0.0001% (n=6). After testing for fertilization compatibility, a fixed ratio of sea urchin eggs and activated sperm suspension was added to each vial. The test was stopped after 15 minutes by adding two drops of paraformaldehyde (4% in FSW). Eggs were counted and fertilization success determined in each sample under stereomicroscope by two different observers to increase accuracy. Fertilization success was determined by the formation of thefertilization membrane around the egg cell (Figure A-9). This experiment was repeated three times with different combinations of egg and sperm donors.

Data Analysis

Differences in larval deformation (% deformed of total larvae) and fertilization success (% fertilized eggs of total) among scrubber water treatment concentrations were examined by 1-factor PERMANOVAs (square root transformed data, Euclidian distance matrices) using the Primer-E (6.1.13) software with the PERMANOVA+ (1.0.3) extension (Clarke and Gorley 2006). All PERMANOVA tests were preceded by PERMDISP tests to verify homogeneity of dispersions and followed by pair-wise tests among treatment concentrations. All test results were judged significant using a significance level of 0.05. The lowest treatment concentration showing a statistically significant difference from the control treatment signified the lowest observed effect concentration (LOEC). The next concentration below theLOEC were signified as the no observed effect concentration

(NOEC).The EC_{10} and EC_{50} values for larval deformation and fertilization success were obtained by applying the Excel Macro REGTOX, according to Vindimian et al. (1983). The REGTOX models were based on the equation by Hill.

Results

Development of sea urchin larvae

The percentage of deformed larvae differed significantly among scrubber water treatment concentrations (1-factor PERMANOVA: Pseudo-$F_{8,62}$ = 14.072, p= 0.0001), with the percentage of deformed larvae increasing from 23% ± 10 (SD) in the controls to 97% ± 6.6 (SD) at the highest tested scrubber water concentration 10% (Figure A-8, Figure A-9 left). It was, however, apparent already at much lower concentrations that larvae did not develop in thesame way as they did in the controls. LOEC was determined at a scrubber water concentration of 0.1% (PERMANOVA pair-wise tests: p = 0.021) and NOEC at the next lower test concentration 0.01%. The EC10 and EC50 values were estimated at scrubber water concentration of 2.677% (95% CI: 1.642, 4.303) and 4.678% (95% CI: 3.6, 7.505) respectively (Figure A-9 right). Combustion particles were observed both in the stomachs of feeding larvae and attached to decaying or dead larvae (e.g., Figure A-8, 5%).

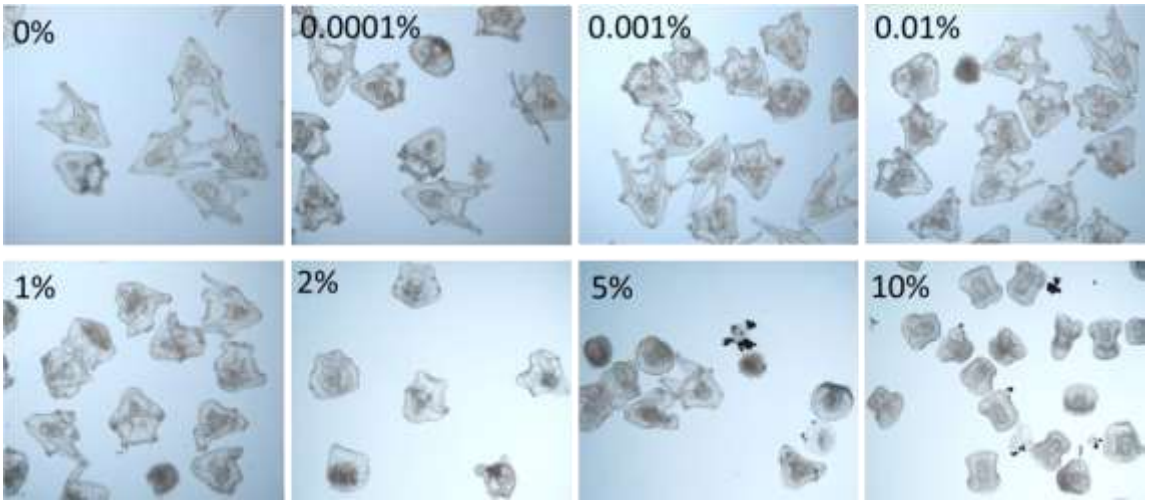


Figure A-8. Larval morphology on day 10 in the exposures of different scrubber water dilutions, from 0% (control) to 10% scrubber water. In the 5% image, combustion particles from the scrubber waterare clearly visible attaching to the biofilm on dead, decaying or undeveloped larvae.

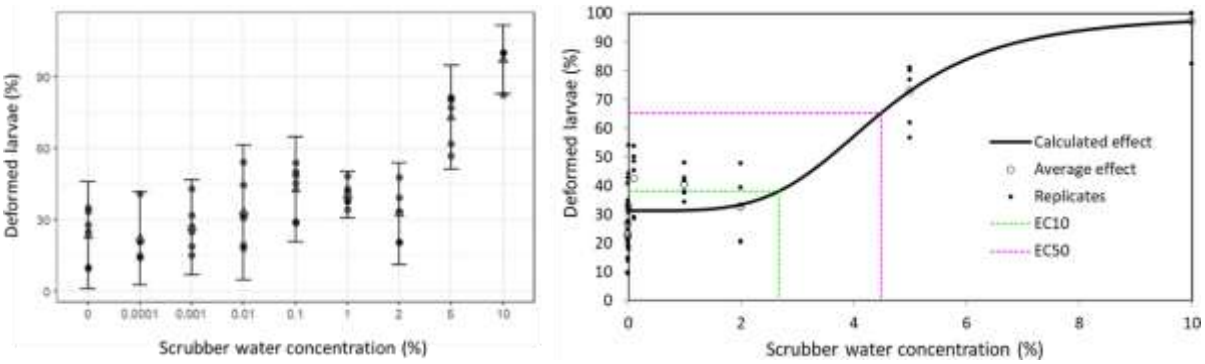


Figure A-9. Deformed larvae (%) at different concentrations of scrubber water (%). Left: average ± SD (n=6 of three pooled timepoints d 8, 9 and 10). Right: Curve fit, and EC10/50 estimation using the REGTOX macro for excel based on the equation by Hill.

Fertilization success of sea urchin eggs

The percentage of unfertilized eggs differed significantly among scrubber water treatment concentrations (1-factor PERMANOVA: Pseudo-$F_{4,53}$ = 7.989, p= 0.0002), with the percentage of fertilized eggs decreasing from 97.6 % ± 1.2 (SD) in the controls to 1.2 % ± 0.87 (SD) at the highest tested scrubber water concentration 50% (Figure A-10, Figure A-11 left). LOEC was determined at the lowest tested scrubber water concentration of 0.0001% (PERMANOVA pair-wise tests: p= 0.0015) and NOEC could thus not be determined other than being <0.0001%. The EC10 and EC50 values were estimated at scrubber water concentration of 2.339% (95% CI: 1.847,2.860) and 7.708% (95% CI: 7.053, 8.463) respectively (Figure A-11 right).

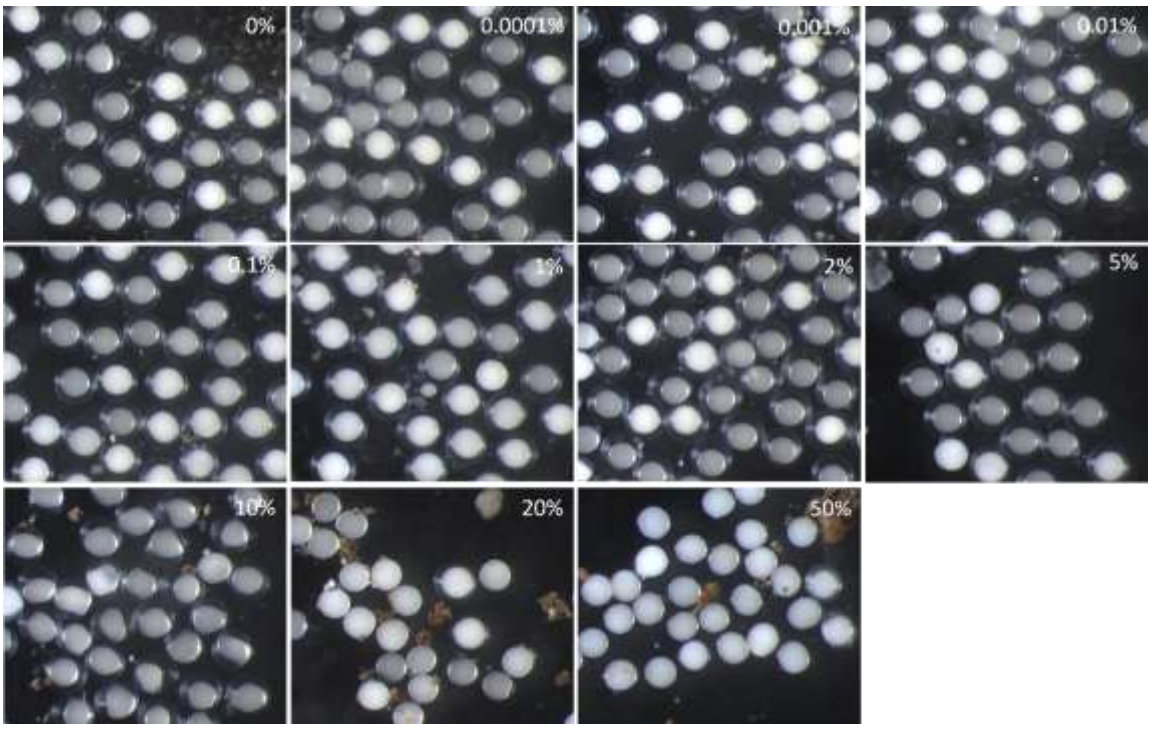


Figure A-10. Eggs of the green sea urchin (*Strongylocentrotus droebachiensis*) after fertilization attempt exposed to different concentrations of scrubber water, from 0% (control) to 50%. Eggs without a fertilization membrane (halo around the egg cell) have failed to become fertilized.

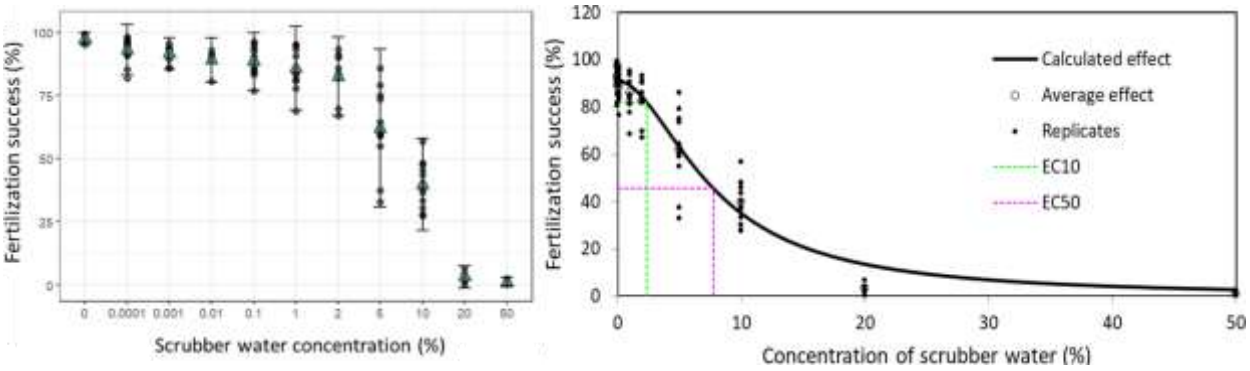


Figure A-11. Fertilization success (%) at different concentrations of scrubber water (%). Left: average ±SD (n=12, two pooled experiments). Right: Curve fit, and $EC_{10/50}$ estimations using the REGTOX macro for excel based on the equation by Hill.

Discussion and conclusions

In all treatments, including the controls, there was a certain number of sea urchin eggs that did not become fertilized, and larvae that did not develop normally. However, in the presence of only 0.0001% scrubber water in the exposure water (1 µL scrubber water per liter), the fertilization was significantly reduced compared to the control, and in a concentration of 0.1% scrubber water the larval development was significantly disrupted

asshown by an increased proportion of malformed larvae (Figure A-10).

Aquatic invertebrates produce very large numbers of eggs and only a fraction of these will develop into adult reproductive individuals even in a pristine, non-polluted area. However, scrubber water from ships is found to have significant effects at such low concentrations that effects on a population level in areas around ship lanes cannot be excluded.

The scrubber water contains numerous compounds and particles known to be harmful to living organisms, e.g., a range of oil-related compounds, elevated concentrations of several metals, and combustion particles like soot, ash, and sulphur particles. It is beyond the scope of the present study to hypothesize which pollutants were responsible for the observed effects, but most likely it was a combination of several of these constituents. The presence ofcombustion particles was obvious when looking at the exposed eggs and larvae in a microscope (see Figure A-8 in exposures to 5 and 10% scrubber water). It was very clear that these particles had a great affinity for organic debris such as dead sea urchin larvae and it seems likely that this could be an efficient transfer route for the particles to the sea floor.

Description on where CRED was NOT followed and why

The ecotoxicological experiments were carried out according to Good Laboratory Practice (GLP) and scored very high in the CRED test. The focus was on ecological relevance where the CRED score was 100%. The reliability score was pulled down by the fact that no standardised method was used. This is, however, the case for all individual studies since standardised tests require regulation of pH as well as filtration of the scrubber water, which would severely affect the relevance of the test results.

References

Clarke, K. R., and R. N. Gorley. 2006. PRIMER v6: UserManual/Tutorial. PRIMER-E, Plymouth.

Koski, M., C. Stedmon, and S. Trapp. 2017. Ecological effects of scrubber water discharge on coastal plankton: Potential synergistic effects of contaminants reduce survival and feeding of the copepod *Acartia tonsa*. Marine environmental research **129**:374-385.

Vindimian, E., C. Robaut, and G. Fillion. 1983. A method for cooperative or noncooperative binding studies using nonlinear regression analysis on a microcomputer. Journal of applied biochemistry **5**:261-268.

Appendix 4: Ecotoxicological results from UAV

Ana Ré & Nelson Abrantes

In the present report we describe the chronic toxicity tests performed with the sea-urchin *Paracentrotus lividus*(fertilization and larval development bioassay) and with the polichaeta *Sabellaria alveolata* (larval development bioassay) exposed to three different scrubber-waters. Tests were always carried out with and without pH correction.

Scrubber water used

Toxicity tests were performed on three different scrubber water samples.

1º The first sample was obtained from the Chalmers University of Technology produced with Atlantic seawater;

2º The second scrubber-water sample was obtained from Catherine C produced with Atlantic seawater.

3º The third scrubber-water was kindly provided by IVL Swedish Environmental Institute produced at Leo Cwith North Atlantic Seawater (ID point 1B_SCRW).

The scrubber water from LEO C collected for the IVL case study was relatively enriched in PAHs and alkylated-PAHs, but for the trace metals, the flask was broken, and we only have the value for Hg (see Table A-1). The artificial scrubber water obtained from Chalmers was characterized by high concentrations of Cr, Mn, Fe, Co, Ni and Cu (Table A-1) while the values of metals in the scrubber-water from Catherine C were higher for V, Zn, As, Cd and U (Table A-1). The chemical analysis performed in an unfiltered sample of Atlantic water revealed high values of Hg and Pb. However, since the water used in the bioassays was filtered, and no negative responses were observed in the control, we can assume that the filtration process led to a reduction in the concentrationof these metals.

Table A–1. Physico-chemical parameters of the seawater dilution and scrubber-water for the differentbioassays.

Physico-chemical parameters	2021 - Preliminary bioassays			2022 - Bioassays				
	Atlantic seawater	Chalmers	Catherine C	Atlantic seawater	Chalmers	Catherine C	Atlantic seawater	Leo C
pH	8.11	2.75 (8.11)	3.34 (8.10)	8.2	2.88 (8.22)	3.37 (8.21)	8.24	2.86 (8.21)
Salinity	35.6	35.5	35.6	35.1	36.1	35.8	34.6	34.6
Dissolved Oxygen	98.8	48.18[(1)]	94.1	99.8	99.7	99.6	99.6	99.7
Temperature	20	19.8	19.2	20	19.9	19.8	21	19.8
Site of sample	N 40º 37,631’	_	N 38º 39,5’	N 40º 37,631’	_	N 38º 39,5’		N 40º 45,403’

Collection	W08º 50,361'		W10º 06,2'	W08º 50,361'		W10º 06,2'		W 08º 50,361'

Methods

The three scrubber-waters were tested using two different species, the sea-urchin *Paracentrotus lividus* and the polychaeta *Sabellaria alveolata* and two chronic toxicity tests (fertilization success and larval development)with and without pH correction:

1) Fertilization success test with the sea-urchin *Paracentrotus lividus*.

The sea-urchin sperm was exposed during 20min to the different scrubber-waters. After, the oocytes were added and exposed during 20min being the assay finished with the addition of formalin. The fertilized eggs are recognized by their fertilization ring as opposed to the unfertilized egg which does not have this characteristic. The test was performed according to the EC(2011) and Beiras et al. (2012), at 20°C using Atlantic seawater (35,1‰) as dilution water. The eight scrubber-water concentrations used were 0.0; 0.01; 0.1; 1.0; 10.0; 25.0; 50.0 and 100-0% with five replicate each.

2) Larval development test with the sea-urchin *Paracentrotus lividus.*

After 48h of exposure at 20ºC the different larval stages of *P. lividus* were evaluated in a total of 100, with the Pluteus larvae being considered the larvae with normal growth and trochophore larvae, gastrulae, blastulae, morula, or split egg as larvae with abnormal development. The test was performed according to EC (2014), Manzo (2010) and Quintino et al. (2009). The test used the same Atlantic seawater as dilution water. The following nine scrubber water concentrations were tested (%): 0.0; 0.001;0.01; 0.1; 1.0; 10.0; 25.0; 50.0 and 100.0 also with five replicate each.

3) Larval development test with the polychaeta *Sabellaria alveolate*.

The polychaeta was exposed during 72h at 20ºC to the different scrubber-waters. The percentage of larvae with long cilia as normal vs. trochophore larvae, gastrulae, blastulae, morula, or split egg was assessed in a total of 100 larvae. The test was performed according to Quintino et al. (2008) and Ré et al. (2007). This bioassay used the same dilution water and the same scrubber-water concentrations as described above for the sea-urchin larval development assay.

Scrubber water samples were tested with and without pH correction. The larval development test with *P. lividus* and *S. alveolata* is an adaptation of the protocol described for *Mytilus*

galloprovincialis in ISO 17244 (ISO, 2015).

Results

Chalmers scrubber water

The scrubber water obtained from Chalmers produced sublethal negative effects on fertilization and larval development at the lowest tested concentrations and without pH correction (NOEC<0,01% on the fertilization bioassay and NOEC=0.001% for the larval development bioassay, both for the test with the sea urchin as well as with the polychaeta) (Table A-2). The polychaeta showed more sensitivity to this scrubber-water with lower EC_{50} and EC_{10} values (Table A-3). When the pH is corrected the toxicity of the scrubber water decrease significantly as would be expected by the change in the availability of some metals. The greater sensitivity of the sea urchin observed in the preliminary bioassays was confirmed by the assays performed later.

Catherine C scrubber water

The scrubber water obtained from DANAOS – Catherine C produced sublethal negative effects on fertilization and larval development at the lowest tested concentrations and without pH correction (see Tables A-2 and A-3). By comparing the results of the larval development assay of *P. lividus* and *S. alveolata*, the sea-urchin revealed high sensitivity with lower NOEC, LOEC, EC_{10} and EC_{50} values, either in the test with or without pH correction. For tests with pH adjustment, EC_{50} and EC_{10} values as well as NOEC and LOEC increase, attenuating the effects observed without pH adjustment.

Leo C scrubber water

The scrubber water obtained from DANAOS – Leo C induced sublethal negative effects on the fertilization and larval development at the lowest tested concentrations and without pH correction (see Tables A-2 and A-3).The Fertilization test was least sensitive compared to the larval development tests to depict the toxicity of scrubber-water, with high toxicological values. By comparing the distinct assays performed, the larval development assay with *P. lividus* showed to be the more sensitive. Once again, the assays with pH correction produced a lower toxicity in all assays performed.

Table A–2. Summary table of NOEC and LOEC values for the bioassays performed with the sea urchin *Paracentrotus lividus* and the polychaetae *Sabellaria alveolata*) exposed to three scrubber waters (Chalmers,Catherine C and Leo C). LOEC – Lowest observed effect concentration, NOEC – No observed effect concentration.

		Chalmers		Catherine C		Leo C	
		Without pH correction	With pH correction	Without pH correction	With pH correction	Without pH correction	With pH correction
Preliminary assays2021	Fertilization	NOEC = 1.56%	NOEC =6.25%	NOEC = 1.56%	NOEC = 1.56%		
	bioassays	LOEC=3.125	LOEC=12.5 %	LOEC=3.125%	LOEC=3.125%	*Not performed*	*Not performed*
	(*P. lividus*)						
	Larval development bioassays (*P. lividus*)	NOEC < 1.56% LOEC=1.56 %	NOEC < 1.56% LOEC=1.56 %	NOEC < 1.56%	NOEC < 1.56%		
				LOEC=1.56%	LOEC=1.56%	*Not performed*	*Not performed*
Assays 2022	Fertilization	NOEC < 0.01%	NOEC =0.10%	NOEC <0.01%	NOEC =0.10%	NOEC =0.01%	NOEC=1.00%
	bioassays						
		LOEC=0.01	LOEC=1.00 %	LOEC=0.01%	LOEC=1.00%	LOEC=0.1%	LOEC=10.0 %
	(*P. lividus*)						
	Larval development bioassays (*P. lividus*)	NOEC =0.001%	NOEC =0.100%	NOEC <0.001%	NOEC =0.010%	NOEC =0.010%	NOEC=0.100 %
		LOEC=0.010	LOEC=1.000 %	LOEC=0.001%	LOEC=0.100%	LOEC=0.100 %	LOEC=1.000 %
	Larval development bioassays (*S. alveolata*)	NOEC =0.001% LOEC=0.01 %	NOEC < 0.01% LOEC=0.01 %	NOEC =0.001%	NOEC < 0.01%	NOEC <0.001 %	NOEC < 0.01%
				LOEC=0.01%	LOEC=0.01%	LOEC=0.001 %	LOEC=0.01 %

Table A–3. Summary table of EC_{10} and EC_{50} values for the bioassays performed with the sea urchin *Paracentrotus lividus* and the polychaetae *Sabellaria alveolata*) exposed to three scrubber waters (Chalmers, Catherine C and Leo C). EC_{10} - 10% Effect concentration, EC_{50} - 50% Effect concentration.

		Chalmers		Catherine C		Leo C	
		Without pH correction	*With pH correction*	*Without pH correction*	*With pH correction*	*Without pH correction*	*With pH correction*
Preliminary assays2021	Fertilization	EC50 = 13.7	EC50 = 81.6	EC50 = 22.9	EC50 = 59.4		
	bioassays	[23.23;14.22]	[79.43;83.78]	[21.61;24.22]	[56.57;62.34]	*Not performed*	*Not performed*
	(P. lividus)						
	Larval	EC50 = 1.3	EC50 = 1.2	EC50 = 1.5	EC50 = 1.5		
	development	[1.22;1.32]	[1.09;1.31]	[1.39;1.62]	[1.35;1.66]		

	bioassays					***Not performed***	***Not performed***
	(P. lividus)						
Assays2022	***Fertilization bioassays (P. lividus)***	**EC50 = 26.68 [23.691;29.675] EC10 = 6.36 [4.560;8.152]**	**EC50 >100 EC10 = 5.233 [2.280;8.185]**	**EC50 = 33.66 [26.775;40.536] EC10 = 7.56 [3.577;11.545]**	**EC50 = 53.43 [48.292;58.563] EC10 = 22.02 [16.972;27.066]**	**EC50 = 11.38**	**EC50 = 58.81**
						[11.028;11.730]	**[48.292;58.563]**
						EC10 = 7.22	**EC10 = 12.96**
						[6.712;7.737]	**[11.874;14.043]**
	Larval development bioassays (P. lividus)	**EC50 = 8.04**	**EC50 = 45.23**	**EC50 = 6.13**	**EC50 = 25.68**	**EC50 = 5.51**	**EC50 = 44.76**
		[5.664;10.413]	**[41.972;48.490]**	**[3.493;8.765]**	**[23.666;27.700]**	**[3.651;7.369]**	**[43.226;46.296]**
		EC10 = 0.265	**EC10 = 20.20**	**EC10 nd**	**EC10 = 7.19**	**EC10 = 0.78**	**EC10 = 24.92**
		[0.057;0.473]	**[16.743;23.653]**		**[5.811;8.576]**	**[0.194;1.362]**	**[22.692;27.156]**
	Larval development bioassays (P. lividus)	**EC50 = 3.80**	**EC50 = 35.27**	**EC50 = 9.44**	**EC50 = 28.16**	**EC50 = 10.47**	**EC50 = 73.29**
		[2.562;5.042]	**[28.634;41.915]**	**[7.478;11.400]**	**[25.727;30.590]**	**[8.362;12.572]**	**[63.873;82.697]**
		EC10 <0.001	**EC10 = 9.18**	**EC10 = nd**	**EC10 = 8.35**	**EC10 = 1.13**	**EC10 = 12.18**
	(S. alveolata)		**[4.829;13.530]**		**[6.587;10.113]**	**[0.533;1.735]**	**[8.091;16.262]**

Discussion and conclusion

All scrubber waters tested showed an enrichment of PAHs and alkylated-PAHs, as well as trace metals, all of them known by their inherent toxicity and persistence. By comparing the concentration of PAHs and alkylated-PAHs the magnitude order found was: Leo C > Catherine C > Chalmers. Concerning the trace metal levels in thescrubber waters, they varied according to the element: V, Zn, As, Cd and U were found higher at Catherine C while Cr, Mn, Fe, Co, Ni, Cu, Hg and Pb at Chalmers. Leo C was not analysed for metals due to an accident withthe sample.

The exposure of *P. lividus* gametes and *P. lividus* and *S. alveolata* embryos to scrubber-waters evidenced concentration-dependent effects on the explored endpoints, namely fertilization success and larval development abnormality. Effects on larval stages of sea-urchin and polychaete occurred at scrubber-water concentrations sensibly lower, with larval development being most sensitive than fertilization endpoint.

By comparing the three tested scrubber-waters, the one that was more toxic for the sea urchin, either for thefertilization or the larval development, was the scrubber-water from Leo C, which corresponds to the scrubber- water that presents the higher 16-PAHs and Alkyl PAHs concentrations. Concerning the larval development ofthe polychaeta, the most toxic scrubber

water was Chalmers.

All the assays performed with pH correction, as it is mandatory to perform in scrubber water before their release, showed lower toxicity, which in part can be explained by the decrease in metal's bioavailability.

References

Beiras, R., Durán, I., Bellas, J., and Sánchez-Marín, P. 2012. Biological effects of contaminants: *Paracentrotuslividus* sea urchin embryo test with marine sediment elutriates. ICES Techniques in Marine Environmental Sciences No. 51. 13 pp.

Environment Canada 2011. Biological Test Method: Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars) EPS 1/RM/27 Second Edition – February 2011 Science and Technology Branch Environment Canada

Environment Canada 2014. Reference Method for Measuring the Toxicity of Contaminated Sediment toEmbryos and Larvae of Echinoids (Sea Urchins or Sand Dollars) 1/RM/58 July 2014. Science and Technology Branch Environment Canada

ISO, 2015. Water qualiy — Determination of the toxicity of water samples on the embryo-larval development of Japanese oyster (*Crassostrea gigas*) and mussel (*Mytilus edulis* or *Mytilus galloprovincialis*). ISO 17244:2015.

Quintino V., Rodrigues A.M., Freitas R., **Ré A.** (2008) Experimental biological effects assessment associatedwith on-shore brine discharge from the creation of gas storage caverns. *Estuarine Coastal and Shelf Science. 79, 525-532.* http://doi:10.1016/j.ecss.2008.05.004

Ré, A., Rodrigues, A. M., Quintino, V. (2007) Acute toxicity testing with the European estuarine amphipodCorophium multisetosum. *Hydrobiologia, 587,* 89-99.

Silva, S., **A. Ré**; P. Pestana, A.M. Rodrigues & V. Quintino (2004). "Sediment disturbance off Tagus Estuary,

Western Portugal: chronic contamination, sewage outfall operation and runoff events" *Marine PollutionBulletin* 49, 154-162. http://doi: 10.1006 / j.marpolbul.2004.02.004

Appendix 5: Ecotoxicological results from AUTH

Savvas Genitsaris, Polyxeni Kourkoutmani, Natassa Stefanidou & Maria Moustaka-Gouni

Scrubber water used

The scrubber effluent that was used during the experimental work was collected according tothe standardized protocol from ICRA. The effluent was sampled from an open-loop system onboard the DANAOS Leo C container ship at the Mediterranean Sea Site 11 (11B_SCRW) on the 22nd of November 2021. PAHs, alkylated PAHs, and metals concentrations were at similar levels for several compounds compared to 10B_SCRW. Acenaphthene, Pyrene, Dibenzo(a,h)anthracene, Benzo(g,h,i)perylene, Naphthalene-2-methyl, Naphthalene-C2, Naphthalene-C3, V, Mn, Cu and Zn were lower compared to 10B_SCRW. But Naphthalene (only Filtrate) was much higher than that of 10B_SCRW.

Methods

1) Experimental design and conditions

We performed four sets of ecotoxicological experiments using mesocosms as experimental units. Twelve glass mesocosms of 12 L were set up indoors in temperature-controlled environments for each experiment. The natural plankton communities of Thermaikos Gulf (Port of Thessaloniki, Th) and Plagia (P) were subjected to two scrubber effluent dilution (10 and 1 % v/v) treatments (indicated as HS and LS, respectively) while the natural plankton communities of Saronikos Gulf (Flisvos Marina, F) and Vouliagmeni (V) were subjected to three scrubber effluent dilutions (1, 2, and 5 % v/v, indicated as LS, MS and HS, respectively). In all four experiments one control (C) was set up in conjunction with one control enriched in nitrate -nitrogen (HN) mimicking the -N concentration of 10% treatment only in the case studies of Th and P. For each treatment three replicates were used. During the experiments temperature and salinity levels were controlled within the range of the in-situ conditions and pH ranged between 7.5 in high scrubber treatments (10 %) and 8.1 in controls in Thermaikos experiments, and between 7.3 in high scrubber treatments (5 %) and 8.1 in controls in Saronikos experiments. The ecotox experiments were terminated on day 6 due to the formation of the first aggregates and the inspection for mesocosm wall-growth effect to avoid artifacts.

Mesocosms were used as a valuable tool to fill the scale gap between laboratory experiments based on single species / clonal cultures and field studies. Thus, the experimental conditions in treatments will select for the best adapted genotypes within a highly diverse starting community instead of selecting from one species or the extremely restricted variance of communities

assembled from stock cultures (Moustaka-Gouni et al. 2016). Although mesocosms are a well-developed method in both aquatic ecology and ecotoxicology (e.g., used to assess the environmental risks posed by pesticides) it is not considered a toxicological standardized method. Nevertheless, mesocosms can be a valuable tool to assess the impact of a chemical on populations or communities of aquatic ecosystems under more realistic environmental conditions than is achievable with standard single-species laboratory studies (EC 2018). The following criteria that should be addressed when assessing mesocosm data according to EC TGD 27 (2018) were fulfilled in our case study:

- the experimental set-up of mesocosms is adequate for assessing the effects of scrubber effluent on plankton communities
- realistic communities were used (the natural phytoplankton and bacterioplankton communities) and the species composition in the mesocosms was representative to those in the field
- there was adequate description of exposure patterns; our compartment of interest was the water column containing the most suitable and realistic communities
- sensitive test endpoints that are in accordance with the mode of action of the chemical were implemented. The test endpoints that were selected are considered robust and preferable in regulatory documents while a broad spectrum of species sensitivity was tested. Representatives of diverse taxonomic, size, and functional groups of phytoplankton and bacterioplankton were included.

However, the scientifically robust statistical evaluation of the results could be limited due to the relatively low number of replicates used for each treatment and controls, because of the high amount of labour involved in microscopy analysis (counting more than 240000 plankton individuals) plus the associated molecular analysis. We consider the Mann-Whitney test to be the appropriate statistical approach to test for differences between treatments. To increase the test's sensitivity, we aggregated the abundances of each of the dominant species during the first three days period (thus 9 values), reflecting acute/chronic responses.

2) Test organisms and ecotoxicological test endpoints

Near-surface seawater was collected from all sites containing the winter phytoplankton and bacterioplankton communities. The metazoan grazers were removed by sieving the communities through a 200 μm mesh size gauze. The whole size spectrum (pico-, nano- and micro-) and taxonomic diversity (diatoms, dinoflagellates, cryptophyes, haptophytes, pico-chlorophytes and pico-cyanobacteria) of phytoplankton and bacterioplankton were represented.

The phytoplankton species names of the test organisms and bacterioplankton taxonomic groups and Operational Taxonomic Units (OTUs) have been defined. An appropriate taxon-dependent test duration was defined. For bacteria and microalgae, the ecotoxicological growth tests for 5 to 7 days are considered multigeneration tests. Within thistimeline, the species response was examined every 24 hours, which is appropriate for studyingbacterial and algal growth. Also, 24 hours duration exposure minimizes the complications dueto volatilization and degradation of the contaminants. Based on the guidelines of TGD 27 (EC 2018) the following endpoints have been examined: growth (growth rate) and numbers (population density, total abundance) as well as community composition shift (community structure) although not included in the guidelines for algae and bacteria.

3) Microscopic analysis

The samples were examined using the inverted epifluorescence microscope Nikon ECLIPSE TE2000-S. Counting was done using the inverted microscope method. For phytoplankton with size < 3 μm and for bacterioplankton, subsamples of 10 ml were fixed with formaldehyde, incubated for at least 24 h (at 4 °C in the dark), filtered onto black Nuclepore filters (0.2 μm pore size) and stained with DAPI. The filters were observed under ultraviolet, green, and blue excitations at 1000X.

4) Metabarcoding and metagenomic high-throughput sequencing and bioinformatic pipelines

Molecular analyses were performed in the experiments of Thermaikos Gulf. For this purpose, subsamples of 500 mL were filtered through 0.2 μm nucleopore filters for metabarcoding high-throughput sequencing (HTS) on Day 0, and on Days 2 and 4 from the control and scrubber treatments. DNA was extracted using a Macherey-Nagel NucleoSpin® Soil, GenomicDNA Isolation Kit, according to the manufacturer's instructions. The concentration and qualityof recovered DNA was confirmed using the Thermo Scientific™ NanoDrop™ spectrophotometer. For metabarcoding sequencing, the extracted DNA was subjected to PCR using specific primers targeting the V3-V4 hyper variable regions of the 16S rRNA gene (S-DBact-0341-b-S-17 = CCTACGGGNGGCWGCAG; S-D-Bact-0785-a-A-21 = GACTACHVGGGTATCTAATCC) for the bacterioplankton. The amplicon samples were subjectedto High-Throughput Sequencing on Illumina MiSeq using 300 + 300 bp paired-end chemistry. For metagenomic sequencing, DNA extracted from the subsamples of Day 0 and equimolar integrated DNA from all replicates of the HS treatment in Day 2, was used on

Illumina MiSeq using 300 + 300 bp paired-end sequencer. The PCR amplification step and sequencing steps for both metabarcoding and metagenomic analyses were performed at the Mr DNA Molecular Research Laboratory, Shallowater, USA, according to in-house protocols. Bioinformatic pipelines were performed according to standardized methods. Final datasets with clean readswere taxonomically and functionally annotated using appropriate public databases (e.g., the SILVA database) and relevant tools (e.g., BLAST searches).

5) Data analysis

Firstly, all cell abundance data were checked for normal distribution and variance homogeneity with the Shapiro-Wilks test, showing non-normal distribution in all cases. Then the Levene's test was applied to assess the equality of phytoplankton abundance variances among the replicates of the same treatment indicating no statistical differences among replicates of the same treatments, thus average cell abundances were used to compare treatments in subsequent analyses. The Mann-Whitney non-parametric test was applied to evaluate whether the median of the phytoplankton community species abundances as well asthe dominant species abundance median were statistically different between pairs of samples(control vs treatments). All the analyses were run in R 4.2.0 environment using the multcompackage. The populations net growth rate (r) was determined from changes in the populationabundance between the Day of the maximum abundance of the population in the controls (referenced as Critical Time in the rest of the text) and Day 0 divided by the number of days between those two times, according to the equation:

$$r = \frac{lnB_{CT} - lnB_0}{D_{CT} - D_0}[d^{-1}]$$

where B is the cell abundance of the population, CT is the critical time (day) of maximum observed cell abundance of the population in the controls, and D represents the Day of the measurement.

The bacterial assemblages of the different treatments in Days 2 and 4 derived from the natural community of the Port of Thessaloniki (Day 0) were compared with controls using thePlymouth routines in the multivariate ecological research software package (PRIMER v.6). TheBray–Curtis dissimilarity coefficients were calculated to build the matrix based on bacterial OTUs number of reads in order to identify interrelationships between the samples, and a multidimensional scaling plot (nMDS) was constructed.

Results

Phytoplankton (Nano- and micro-plankton eukaryotes)

The cell densities in the treatments of different scrubber dilutions showed a similar pattern for total phytoplankton and dominant species with an immediate and sharp decline in cell numbers at the highest scrubber effluent concentrations in most cases (see Figure A-12 for the experiment of the communities of Thessaloniki Port). A similar pattern was also observed for the growth rate at Critical Time (CT, the time of maximum growth of control) (Figure A-12) with theexceptions of the species *Cylindrotheca closterium* in P, and *Chrysochromulina* sp. in F, showing positive growth rate even at high scrubber treatments (10 % and 5% v/v, respectively)(not shown). The nitrate-nitrogen enrichment in the experiments of Thermaikos Gulf showeda non-significant increase of the total phytoplankton and the dominant diatom in Th, the potentially toxic, bloom-forming *Pseudonitzschia* cf. *pungens*. This increase did not mask the significantly negative, toxic effect of scrubber effluent (10 % v/v) on its growth (Figure A-12).

The cell densities of pico-plankton (calculated every 24-hours) in treatments showed a different pattern compared to that of nano- and micro- phytoplankton eukaryotes. Picophytoplankton was the dominant component in terms of abundance of the phytoplankton communities in the unpolluted studied areas P and V. The cell densities and growth rates in the scrubber treatments showed a similar pattern for P and V indicating no negative effect even at the highest scrubber effluent concentrations for each experiment (Figures A-13─A-14). In the experiments of the polluted areas Th and F, the results are different. In F, picophytoplankton was under the detection limit of epifluorescence microscopy method used while in Th a zero-growth rate was estimated at the highest scrubber effluent concentration (Figure A-15) due to the grazing effect of heterotrophic nanoflagellates (HNFs) (however, no direct negative scrubber effect was detected).

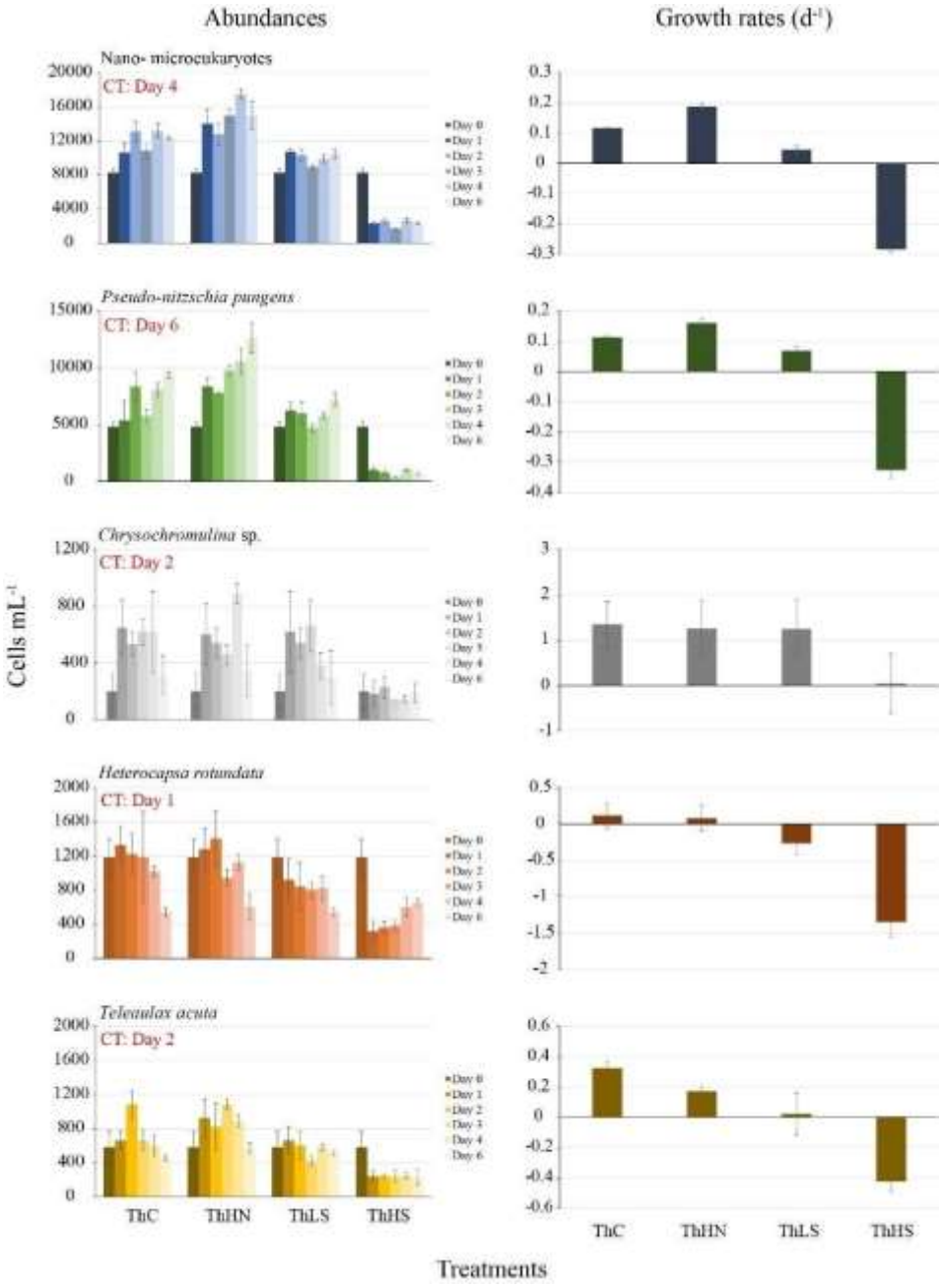


Figure A-12. Left panel: Cell densities every 24-hours in Th mesocosms, in controls and different treatments. Right panel: Growth rates at CT for each dominant species and total nano- and micro-phytoplankton eukaryotes. ThC: Thessaloniki Port control, ThHN: Thessaloniki Port high nitrate control, ThLS: Thessaloniki Port 1 % scrubber, ThHS: Thessaloniki Port 10 % scrubber.
Phytoplankton (pico-plankton)

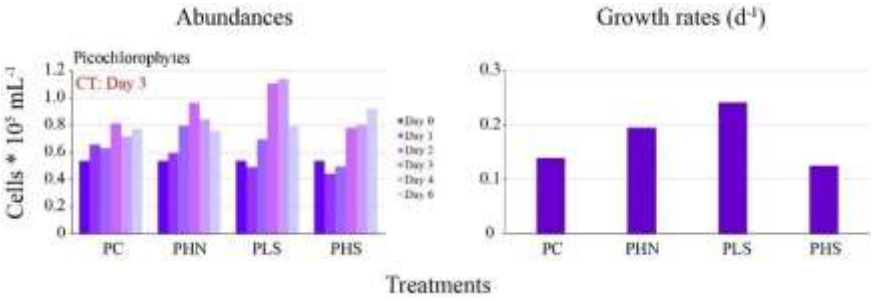


Figure A-13. Left panel: Cell densities every 24-hours in Plagia mesocosms, in controls and different treatments. Right panel: Growth rate at CT for total pico-phytoplankton. PC: Plagia control, PHN: Plagia high nitrate control, PLS: Plagia 1 % scrubber, PHS: Plagia 10 % scrubber.

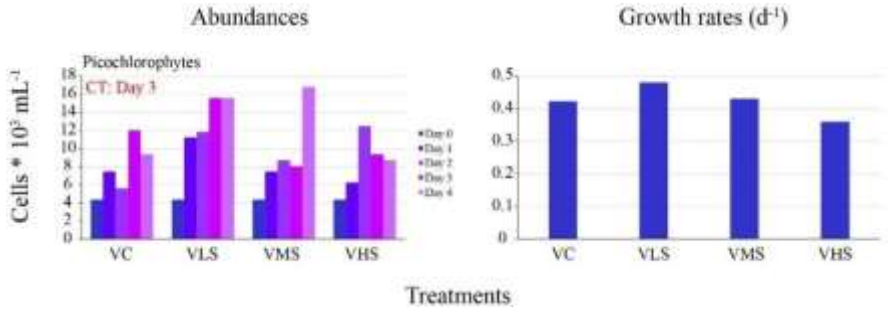


Figure A-14. Left panel: Cell densities every 24-hours in Vouliagmeni mesocosms, in controls and different treatments. Right panel: Growth rate at CT for total pico-phytoplankton. VC: Vouliagmeni control, VLS: Vouliagmeni 1 % scrubber, VMS: Vouliagmeni 2 % scrubber, VHS: Vouliagmeni 5 % scrubber.

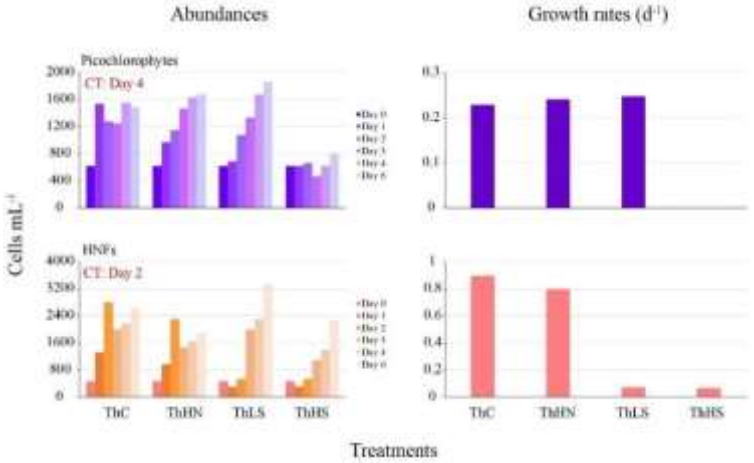


Figure A-15. Left panel: Cell densities every 24-hours in Thessaloniki Port mesocosms, in controls and different treatments. Right panel: Growth rate at CT for total pico-phytoplankton and their grazers HNFs. Th: Thessaloniki Port control, ThHN: Thessaloniki Port high nitrate control,ThLS: Thessaloniki Port 1 % scrubber, ThHS: Thessaloniki Port 10 % scrubber.

Bacterioplankton

The cell densities of bacterioplankton in treatments showed a different pattern compared to that of nano- and micro- phytoplankton eukaryotes but similar to that of picophytoplankton. No negative effect was observed even at the highest scrubber effluent concentrations for each experiment (see Figure A-16 for the experiment of the communities of Flisvos Marina), both in the polluted and unimpacted areas communities. Net growth rates were always positive.

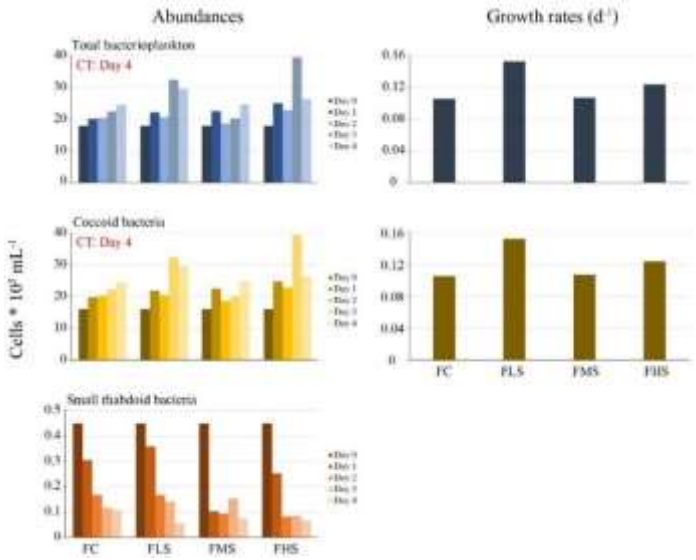


Figure A-16. Left panel: Cell densities every 24-hours in Flisvos Marina mesocosms, in controls and different treatments. Right panel: Growth rate at CT for bacterioplankton. FC: Flisvos control, FLS: Flisvos 1 % scrubber, FMS: Flisvos 2 % scrubber, FHS: Flisvos 5 % scrubber.

Bacterioplankton community structure

The molecular diversity of the bacterial communities examined in Thermaikos Gulf experiments (Th and P) exhibited high richness with over 500 OTUs in each area. The bacterial assemblages of the controls and different treatments in Days 2 and 4 and the Thessaloniki Port natural community (Th Day 0) were compared for dissimilarity of bacterial OTUs numberof reads to identify possible shifts in species composition (community structure) within the bacterioplankton community due to scrubber effluent exposure. The interrelationships between the samples in the multidimensional scaling plot (Figure A-17) shows that bacterial community structure changed from the control to the scrubber (1, 10 %) treatments (3 replicates).

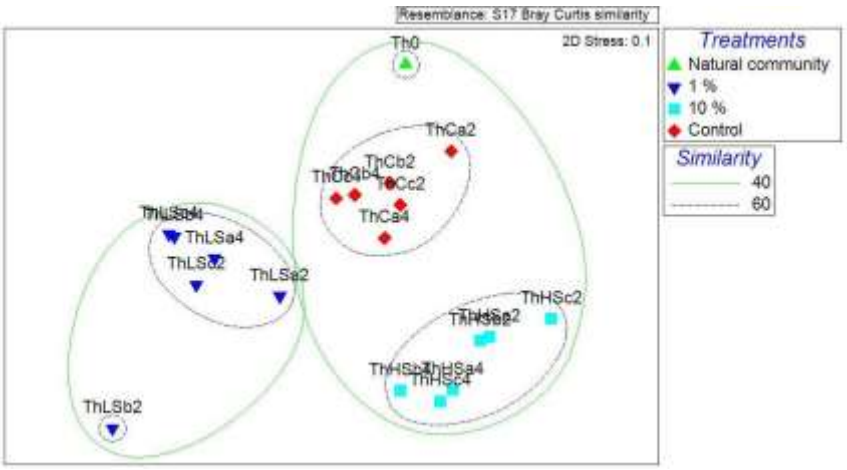


Figure A-17. Multidimensional scaling plot (nMDS) of non-transformed OTUs number of reads in the samples from the Port of Thessaloniki according to Bray-Curtis similarity index. Th0: Initial bacterial community in Day 0. ThC: Control, Day 2, 4, replicates a, b, c. ThLS: Bacterial OTUs 1% scrubber treatment, Day 2, 4, replicates a, b, c, Port of Thessaloniki. ThHS: Bacterial OTUs 10 % scrubber treatment, Day 2,4, replicates a, b, c, Port of Thessaloniki.

NOEC and LOEC

The most sensitive phytoplankton species was the cryptophyte Teleualax sp. in Flisvos marina with LOEC: 2 %. In Voluliagmeni (V), the most sensitive species was again Teleaulax sp. together with the haptophyte species Chrysochromulina sp., and the dinophyte Gymnodinium sp., with LOEC 5 %. In addition to these sensitive phytoplankton species in F experiments the diatom Skeletonema sp. showed a LOEC value of 5 %. In Thermaikos Gulf experiments using the spacing between 1 and 10 % scrubber dilution for most species the LOEC was 10 % in all cases.

Short discussion and Conclusion

The results of the experiments emphasize the need to understand the effects of scrubber effluent on multiple levels and components of the marine environment for higher environmental realism. Attributes at the community and ecosystem level, such as species richness and composition are not demonstrated at lower levels of organization (e.g., at the population level) while they are important test endpoints to assess toxic scrubber effluent effects on marine environment. Nevertheless, such real interactions in natural environmentsmake the results of mesocosm ecotoxicity experiments difficult to be interpreted and thus cannot be used as critical data but only as supporting data according to GD 27 (EC 2018).

Our working hypotheses were:

1. Phytoplankton exposure to scrubber effluent NOx will show a positive response in this eutrophying agent that may mask the short-term adverse effects of PAHs (Ytreberg etal.

2021).

2. Differences in phytoplankton species, taxonomic groups, and size classes sensitivities (pico-, nano-, micro- phytoplankton) to toxicity of PAHs scrubber effluent are expected (Echeveste et al. 2010, Ytreberg et al. 2021)
3. PAHs in scrubber effluent are shaping bacterial communities with indigenous species with biodegradation potential in impacted aquatic environments (Haritash and Kaushik 2009, Ghosal et al. 2016, Hamdan and Salam 2020).

We showed that hypothesis 1 is not supported by the experimental results. Only one phytoplankton species showed positive response to nitrate enrichment that could not mask the severe negative effect of scrubber effluent (10 % treatment). Differences in phytoplankton species and size classes sensitivities (pico-, nano-, micro- phytoplankton) to toxicity of scrubber effluent were observed supporting hypothesis 2. Rich in species (more than 500 OTUs per experiment) bacterial communities with indigenous species with biodegradation potential in aquatic environments were shaped by PAHs of scrubber effluent according to hypothesis 3. Molecular results and statistical analysis show clear evidence for bacterial community shift from the controls to low (1 %) and high scrubber (10 %) treatments in Th community. Indigenous species with biodegradation potential of PAHs were abundant in the scrubber treatments. Similarly, based on preliminary metabarcoding analysis for the experiment in the unimpacted marine community P, several representatives of the PAHs degrading genera in consortia such as *Marinomonas, Mycobacterium, Marinobacter, Halomonas*, *Vibrio, Sulfitobacter, Leisingera, Novosphingobium, Pseudoalteromonas, Polaribacter, Oleispira, Dokdonella, Thalassotalea* occurred in the scrubber treatments. For example, *Marinobacter* that has phenanthrene as a sole carbon source with syntrophic interaction of *Halomonas,* was observed as mixture with *Marinobacter* (Wang et al. 2020). Furthermore, epifluorescence microscopy revealed the long-rod/curved shaped bacterialcellular absorption of PAHs from the medium as bright blue fluorescent blots (Figure A-18). Plankton bacteria had the greatest abundance between all organisms in the seawater exposedto the scrubber water discharge.

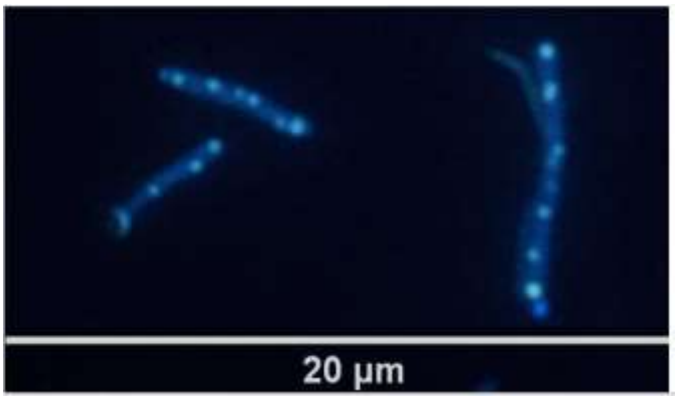


Figure A-18. Micrograph of long rod/curved-shaped bacteria with bright blue fluorescent blots of PAHs in their cells as seen by epifluorescence microscopy. Micrographs were taken under UV excitation for DAPI-staining bacterial cells from a Plagia treatment sample. Scale bar: 20 µm.

In summary, our results revealed that the nano- and micro- algae were negatively affected in experiments at the highest concentrations of scrubber effluent (5 and 10 % dilution in Saronikos and Thermaikos Gulf communities, respectively). Bacterial growth and abundance were not affected negatively at any scrubber effluent concentration in all studied communities. This might be due to bacterial rapid response to scrubber effluent with community shift, dominated by tolerant species and species having degrading capabilities of PAHs contaminants. This may result in natural reduction or even elimination of contaminantsin marine environments and might explain the resistance of the picophytoplankton to the scrubber effluent although it is known as the most sensitive phytoplankon component affected by PAHs (Ben Othman et al. 2023). Plankton algae and bacteria from different areas (Saronikos Gulf polluted and unimpacted sites, Thermaikos Gulf polluted and unimpacted sites) and with different species pools showed repeatable responses to scrubber. The lower impacts of contamination than expected are no surprise, because of the prominent role of ecological interactions of different functional groups in mesocosms and the cometabolic pathways of the scrubber's mixture of PAHs observed in bacterial consortia. Studies that included multiple components of an ecosystem, like ours, were more likely to find no effect of contamination, possibly due to ecological interactions (Johnston et al. 2015).

Description on CRED

The Criteria for Reporting and Evaluating Ecotoxicity Data (CRED) approach is not applicable to multi-species exposure mesocosm studies like our study (Moermond et al. 2016). Nevertheless, to improve the transparency, reproducibility, reliability, and relevance evaluations of EMERGE ecotoxicity studies, we employed the CRED approach where applicable in mesocosm experiments and scrubber effluent as pollutant itself. Based on CRED

analysis high reliability and relevance was evident.

References

European Commission, Directorate-General for Health and Food Safety, Technical guidance for deriving environmental quality standards, European Commission, 2018, https://data.europa.eu/doi/10.2875/018826

Echeveste P, Dachs J, Berrojalbiz N, Agustí S (2010). Decrease in the abundance and viability of oceanic phytoplankton due to trace levels of complex mixtures of organic pollutants. Chemosphere 81: 161–168.

Ghosal D, Ghosh S, Dutta TK and Ahn Y (2016). Current State of Knowledge in Microbial Degradation of Polycyclic Aromatic Hydrocarbons (PAHs): A Review. Front. Microbiol. 7: 1369.

Hamdan HZ, Salam DA (2020). Response of sediment microbial communities to crude oil contamination in marine sediment microbial fuel cells under ferric iron stimulation. Environ. Pollut. 263: 114658.

Haritash AK, Kaushik CP (2009). Biodegradation aspects of polycyclic aromatic hydrocarbons (PAHs): A review. J. Hazard. Mater. 169: 1–15.

Johnston EL, Mayer-Pinto M, Crowe PT (2015). Chemical contaminant effects on marine ecosystem functioning. – J. Appl. Ecol. 52: 140– 149.

Moustaka-Gouni M, Kormas KA, Scotti M, Vardaka E & Sommer U (2016). Warming and acidification effects on planktonic heterotrophic pico- and nanoflagellates in a mesocosm experiment. Protist 167: 389–410.

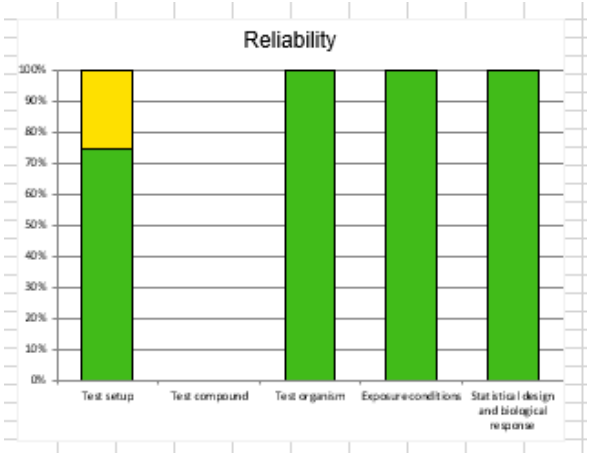
Othman HB, Pick FR, Hlaili AS, Leboulanger C (2023). Effects of polycyclic aromatic hydrocarbons on marine and freshwater microalgae – A review. J. Hazard. Mater. 441: 129869.

Wang C, Huang Y, Zhang Z, Hao H, Wang H (2020). Absence of the nahG-like gene caused thesyntrophic interaction between Marinobacter and other microbes in PAH-degrading process. J. Hazard. Mater. 384: 121387.

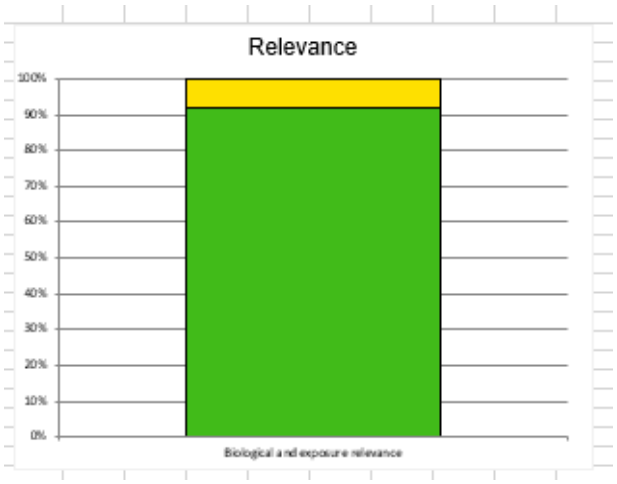
Ytreberg E, Karlberg M, Hassellöv IM, Hedblom M, Nylund AT, Salo K, Imberg H, Turner D, Tripp L, Yong J, Wulff A. (2021). Effects of seawater scrubbing on a microplanktonic community during a summer-bloom in the Baltic Sea. Environ Pollut. 291: 118251.

Appendix 6: CRED analyses UV

Acartia tonsa acute lethality test

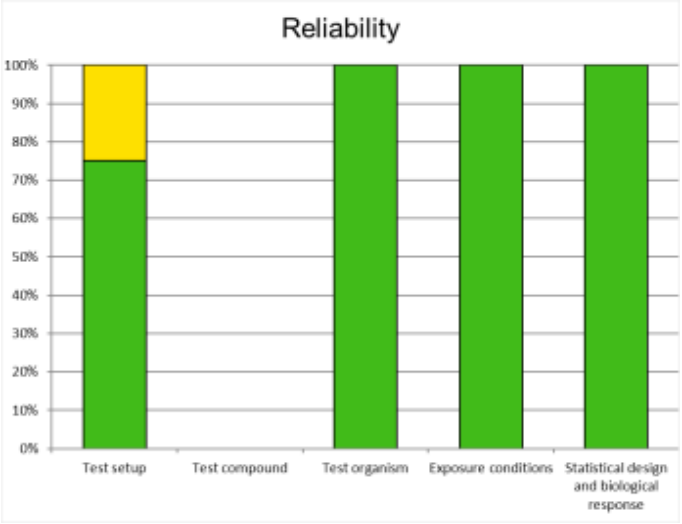


Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	11.11%	
Fulfilled	14	55.56%	
Partially fulfilled	1	27.78%	
Not fulfilled	0	5.56%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Fulfilled	ISO method was followed
1	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Partially fulfilled	We followed the ISO method and a well defined QA/QC program
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	Control survival was checked
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	Negative control and positive control performed (ZnSO4)
	Test compound		
	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Not applicable	
	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Not applicable	
	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not applicable	
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Yes
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	The copepod was obtained from a trustworthy source and then cultured in the laboratory
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	ISO method was followed
1	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	ISO method was followed; conditions werestable during the test.
	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	We used the range 0.01% - 40%
1	14. Is the exposure duration defined?	Fulfilled	24h and 48h
	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Not applicable	
1	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Fulfilled	We followed the ISO recommendations
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	We used 4 replicates with 5 adult copepods each
1	18. Are appropriate statistical methods used?	Fulfilled	We used the one-way ANOVA to calculate the NOEC and the program developed at DTU for calculating Ecxx
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	The curve was observed
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	yes

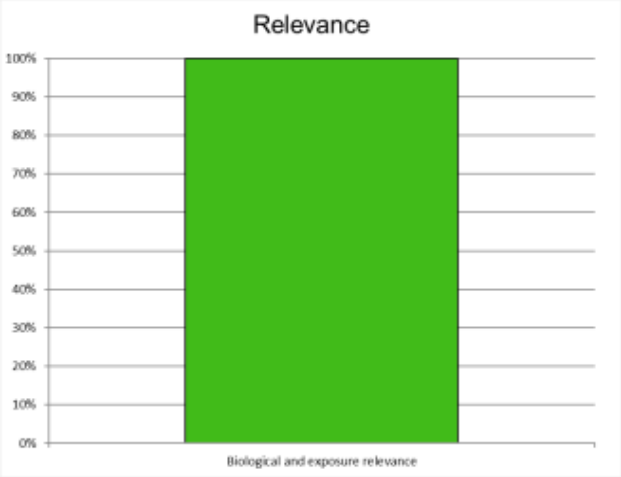


Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	0.00%	
Fulfilled	9	75.00%	
Partially fulfilled	3	25.00%	
Not fulfilled	0	0.00%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Yes, copepods are relevant for the water column
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, copepods are relevant for scrubber water
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Partially fulfilled	This is an acute test, probably not appropriate for regulatory purposes
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	yes
1	6. Are appropriate life stages studied?	Fulfilled	yes
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	yes, NOEC and EC10 were calculated
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	yes
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	
Removed	10. If recovery is studied, is this relevant for the framework for which the study is evaluated?	Not applicable	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	
1	12. Is the tested exposure scenario relevant for the substance?	Fulfilled	
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	

Acartia tonsa larval development test

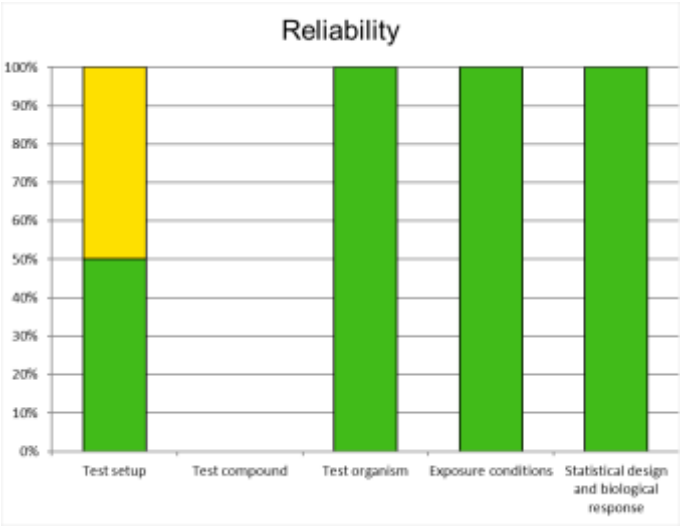


Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	11.11%	
Fulfilled	14	55.56%	
Partially fulfilled	1	27.78%	
Not fulfilled	0	5.56%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Fulfilled	Our testing procedure is very similar to ISO method, but differs from it in some points, including duration (5d in our procedure 5/6d in ISO), acceptability criteria for control (0.5±0.2 after 5d in our procedure; 0.6±0.2 after 5/6d in ISO)
1	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Partially fulfilled	We followed well-defined protocol and QA/QC program
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	Hatching and ratio copepodites/total larvae were checked
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	Negative control and positive control performed
	Test compound		
	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Not applicable	
	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Not applicable	
	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not applicable	
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Yes
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	The copepods were obtained from a trustworthy source and then cultured in the laboratory
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	Yes, experimental system was appropriate and in agreement with ISO standard
1	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	Yes, experimental system was appropriate and in agreement with ISO standard. Conditions were stable during the test
	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	We used the range 0.001% - 40%
1	14. Is the exposure duration defined?	Fulfilled	5 days
	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Not applicable	
1	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Fulfilled	We followed the ISO recommendations
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	We used 6 replicates per scrubber water concentration
1	18. Are appropriate statistical methods used?	Fulfilled	We used the one-way ANOVA to calculate the NOEC and the program developed at DTU for calculating Ecxx
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	The curve was observed
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	yes

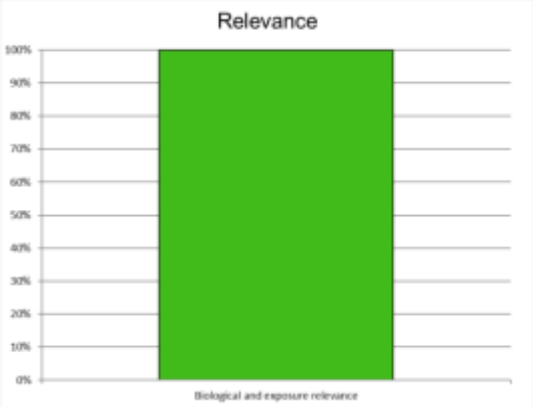


Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	0.00%	
Fulfilled	9	75.00%	
Partially fulfilled	3	25.00%	
Not fulfilled	0	0.00%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Yes, copepods are relevant for the water column
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, copepods are relevant for scrubber water
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Fulfilled	
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	yes
1	6. Are appropriate life stages studied?	Fulfilled	yes
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	yes, NOEC and EC10 were calculated
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	yes
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	
Removed	10. If recovery is studied, is this relevant for the framework for which the study is evaluated?	Not applicable	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	
1	12. Is the tested exposure scenario relevant for the substance?	Fulfilled	
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	

Acartia tonsa long term exposure test

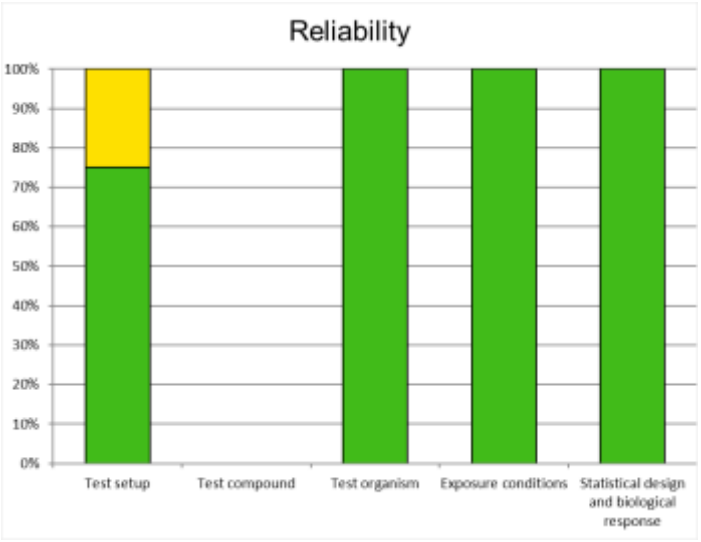


Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	2	11.11%	
Fulfilled	10	55.56%	
Partially fulfilled	5	27.78%	
Not fulfilled	1	5.56%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Partially fulfilled	A standard method for this exposure procedure does not exist. An ISO standard is available for the LDR test, which is one of the tests performed within the long term exposure test (together with F0 egg production)
1	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Partially fulfilled	We followed an internal protocol and a well-defined QA/QC procedure
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	Control survival and ratio copepodite:nauplii was checked in F1 LDR test. Hatching success of F0 was also used as acceptability criterion
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	Negative control and positive control performed (3,5-DCP) for the F_0 generation.
	Test compound		
	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Not applicable	
	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Not applicable	
	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not applicable	
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Yes
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	The copepods were cultured
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	Culturing and LDR test were performed according to ISO standard. Condition was kept constant, and parameters measured at the beginning and end of the LDR test. We did not used aeration to avoid volatilization of PAHs
1	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	Culturing and LDR test were performed according to ISO standard. Condition was kept constant, and parameters measured at the beginning and end of the LDR test. Food provision was accomplished without dilution of the testing concentration.
	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	We used the range 0.001% - 1%
1	14. Is the exposure duration defined?	Fulfilled	yes. Adult copepod at an age of 14 days were used for egg production. Eggs for LDR test were collected after 21 days. LDR test on F1 generation lasted 5 days
	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Not applicable	
1	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Fulfilled	We followed the ISO recommendations
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	We used 10 replicates for the egg production test with F0 generation and 6 replicate for the LDR test on F1 generation. But a single culture was maintained for each scrubber water concentration
1	18. Are appropriate statistical methods used?	Fulfilled	We used the one-way ANOVA to calculate the NOEC and the program developed at DTU for calculating Ecxx
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	The curve was observed
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	yes

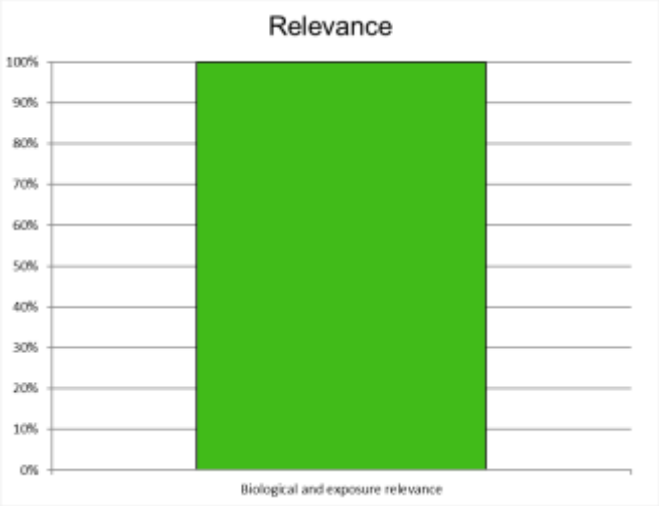


Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	0.00%	
Fulfilled	9	75.00%	
Partially fulfilled	3	25.00%	
Not fulfilled	0	0.00%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Yes, copepods are relevant for the water column
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, copepods are relevant for scrubber water
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Fulfilled	
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	yes
1	6. Are appropriate life stages studied?	Fulfilled	yes
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	yes, NOEC and ECxx were calculated
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	yes
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	
Removed	10. If recovery is studied, is this relevant for the framework for which the study is evaluated?	Not applicable	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	
1	12. Is the tested exposure scenario relevant for the substance?	Fulfilled	
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	

Bivalve larval development test

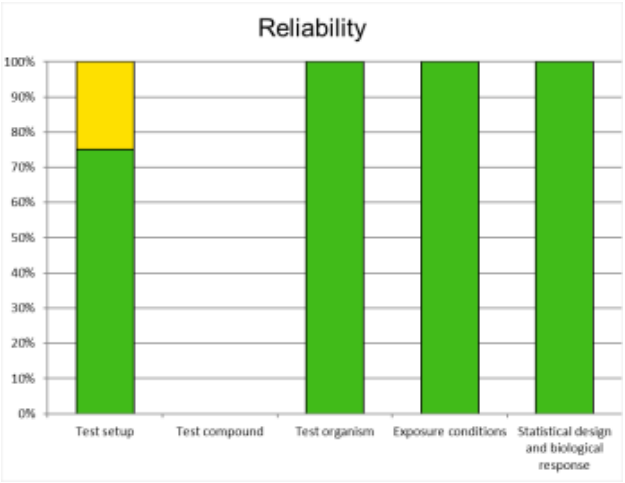


Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	11.11%	
Fulfilled	14	55.56%	
Partially fulfilled	1	27.78%	
Not fulfilled	0	5.56%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Fulfilled	We followed the shared protocol based on the ISO standard
1	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Partially fulfilled	We followed a detailed protocol and a well defined QA/QC program
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	Normal development in control > 80%
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	Negative control and positive control performed
	Test compound		
	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Not applicable	
	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Not applicable	
	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not applicable	
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Yes
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	The bivalve were kept in the lab for at least 5 days before testing
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	Yes, experimental system was appropriate and in agreement with ISO standard
1	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	Yes, experimental system was appropriate and in agreement with ISO standard. Conditions were stable during the test
	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	We used the range 0.001% - 40%
1	14. Is the exposure duration defined?	Fulfilled	48-h
	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Not applicable	
1	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Fulfilled	
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	We used 5 replicates
1	18. Are appropriate statistical methods used?	Fulfilled	We used the one-way ANOVA to calculate theNOEC and the program developed at DTU for calculating Ecxx
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	The curve was observed
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	yes

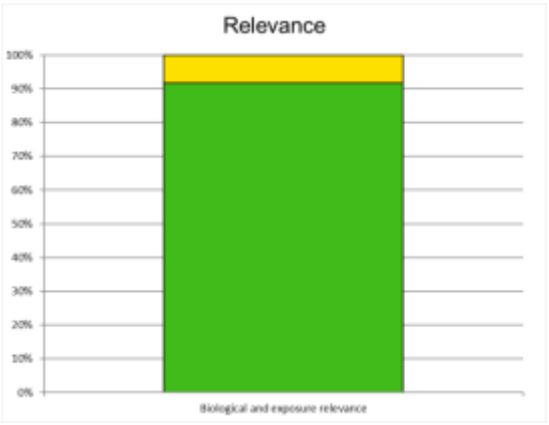


Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	0.00%	
Fulfilled	9	75.00%	
Partially fulfilled	3	25.00%	
Not fulfilled	0	0.00%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Yes, bivalve larvae are relevant for the water column
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, bivalve larvae are relevant for scrubber water
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Fulfilled	
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	yes
1	6. Are appropriate life stages studied?	Fulfilled	yes
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	yes, NOEC and EC10 were calculated
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	yes
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	
Removed	10. If recovery is studied, is this relevant for the framework for which the study is evaluated?	Not applicable	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	
1	12. Is the tested exposure scenario relevant for the substance?	Fulfilled	
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	

Microtox test



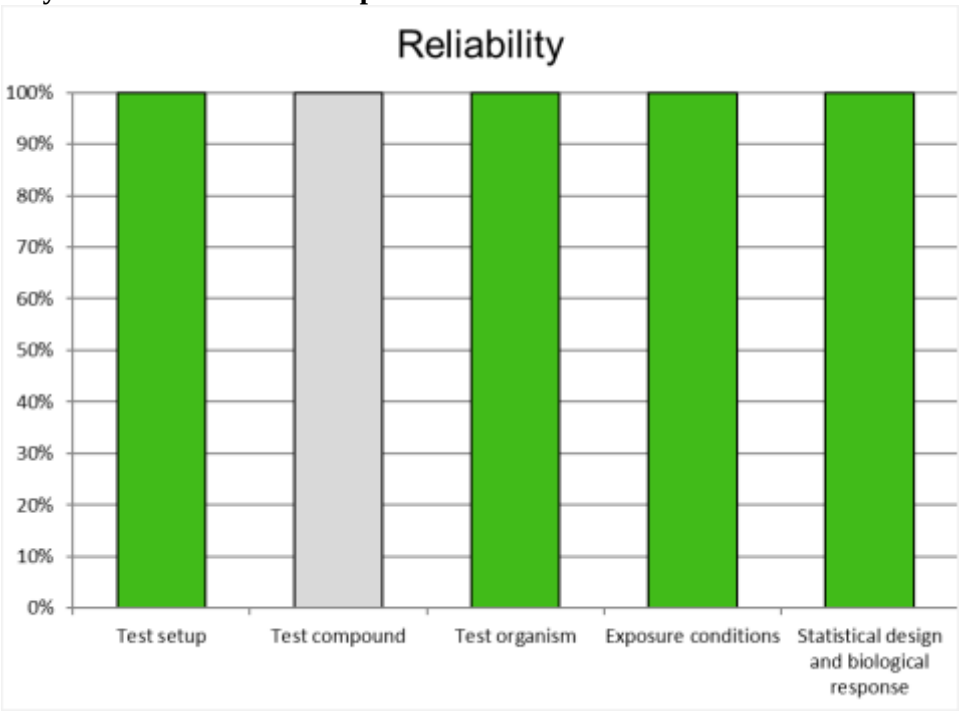
Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	11.11%	
Fulfilled	14	55.56%	
Partially fulfilled	1	27.78%	
Not fulfilled	0	5.56%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Fulfilled	ISO method was followed
1	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Partially fulfilled	We followed the ISO method and an appropriate QA/QC program
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	Bioluminescente at t0
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	Negative control and positive control performed (ZnSO4)
	Test compound		
	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Not applicable	
	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Not applicable	
	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not applicable	
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Liophilised bacteria were used
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	The source of liophilised bacteria is trustworthy. The bacteria were not preexposed to stressors
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	We used the M500 Analyser developed for the Microtox test
1	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	We used the M500 Analyser developed for the Microtox test
	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	We used the range 0.01% - 40%
1	14. Is the exposure duration defined?	Fulfilled	Bioluminescence inhibition was measured after 5, 15 and 30 minutes
	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Not applicable	
1	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Fulfilled	
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	We used the maximum number of replicates allowed by the M50 Analyser
1	18. Are appropriate statistical methods used?	Fulfilled	We used the one-way ANOVA to calculate the NOEC and the program developed at DTU for calculating Ecxx
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	The curve was observed, but the effects are low
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	



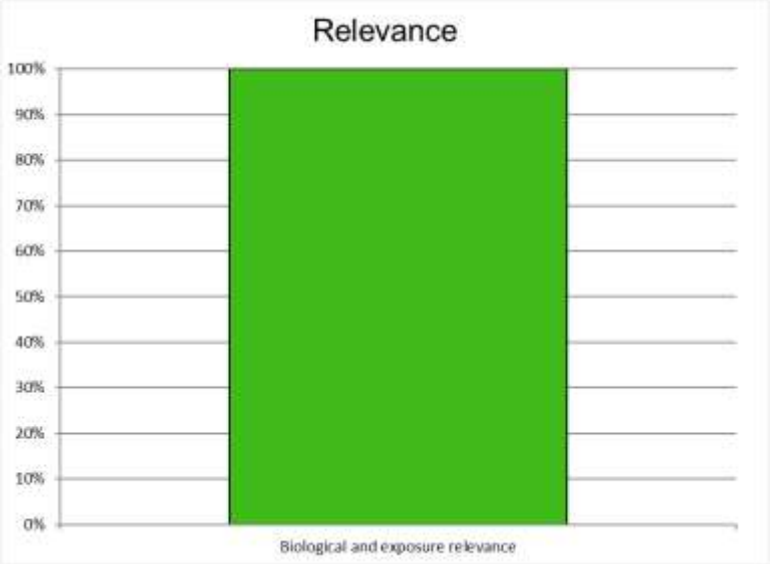
Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	0.00%	
Fulfilled	9	75.00%	
Partially fulfilled	3	25.00%	
Not fulfilled	0	0.00%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Yes, marine bacteria are relevant for the water column
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, marine bacteria are relevant for scrubber water
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Partially fulfilled	This is an acute test, probably not appropriate for regulatory purposes
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	yes
1	6. Are appropriate life stages studied?	Fulfilled	yes
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	yes, NOEC and EC10 were calculated
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	yes
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	
Removed	10. If recovery is studied, is this relevant for the framework for which the study is evaluated?	Not applicable	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	
1	12. Is the tested exposure scenario relevant for the substance?	Fulfilled	
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	

Appendix 7: CRED analysis UoS

Mytilus edulis larval development

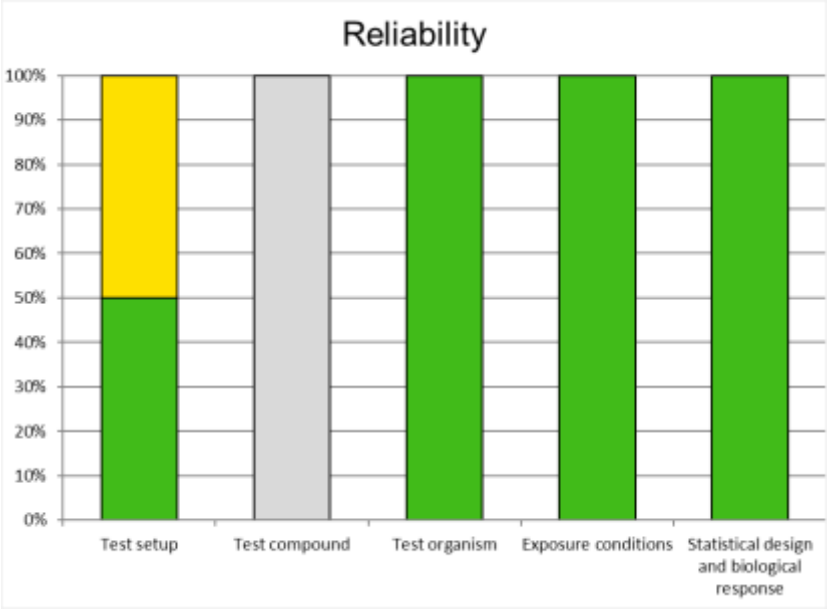


Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	3	15.79%	
Fulfilled	16	84.21%	
Partially fulfilled	0	0.00%	
Not fulfilled	0	0.00%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Fulfilled	Protocol based on the ISO 17244 (2015) Water quality. Determination of the toxicity of water samples on the embryo-larval development of Japanese oyster (*Crassostrea gigas*) and mussel (*Mytilus edulis* or *Mytilus galloprovincialis*). ISO 17244:2015, pp 24
1	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Fulfilled	Methods have been well described and standardized
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	Control survival and normal larvae development
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	Positive and negative controls included
	Test compound		
1	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Not reported	Scrubber water is a mixture of pollutants
1	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Not reported	Scrubber water might be varied under different conditions
1	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not reported	Other compounds could be present and have an effect on organisms
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Yes. *Mytilus edulis*
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	Organisms were collected in field and acclimatized for a week before performing any exposure or treatment. Organisms kept in filtered seawater regularly tested for pollutants and at ambient/tested temperature. Adults were fed at libitum with kelp or macroalga (sea lettuce, genus Ulva). No contaminants added before the experiment and pH
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	Static conditions. Glass crystallization dishes (50 ml) were used as test vessels filled with 20 ml of experimental water. Glass material used in all the steps to avoid the union of the compounds to the walls of the containers.
1	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	Experimental system was appropriate and in agreement with ISO standard. Variables such as temperature, salinity, pH, and DO were measured at the beginning of the experiment. Tests were conducted for 72h at 16ºC and 16h:8h light:darkness photoperiod. Water levels were checked throughout the exposure period, and they were adjusted only when necessary. pH correction was made when necessary, at the beginning of the experiment. No significant pH and OD variation should be observed during the exposure.
Removed	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	Yes, 10 dilutions: 0.001; 0.01; 0.1; 1.0; 2.0; 5.0; 10.0; 20.0; 40.0; 100.0%
1	14. Is the exposure duration defined?	Fulfilled	30min, 72h
1	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Fulfilled	Yes
Removed	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Not applicable	
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	Yes. 5 replicates per treatment. Each container contains enough embryos at the beginning to obtain at least 100 larvae at the end of the exposure (72h)
1	18. Are appropriate statistical methods used?	Fulfilled	LC50 using probit in SPSS. One-way ANOVA and Student's t-test analyses were performed to compare exposure treatments and controls. LOECs and NOECs were determined using one side Dunnett's post hoc tests or the equivalent test for non-parametric analyses
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	Yes
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	Yes



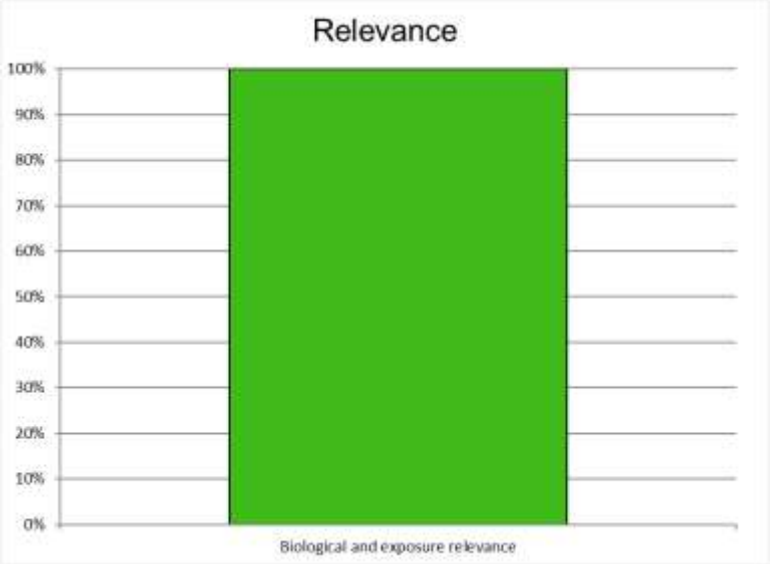
Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	0.00%	
Fulfilled	12	100.00%	
Partially fulfilled	0	0.00%	
Not fulfilled	0	0.00%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Yes, embryo-larval toxicity testing of marine pelagic larvae of *Mytius edulis*
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, marine pelagic larvae of *Mytius edulis*
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Fulfilled	Yes. Fertilization success, larval development
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	
1	6. Are appropriate life stages studied?	Fulfilled	Yes. Embryos and pelagic larvae
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	Yes
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	Yes, seawater
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	Yes, 30 min for fertilization and 72h for larvae
Removed	10. If recovery is studied, is this relevant for the framework for which the study is evaluated?	Not applicable	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	Yes, scrubber effluent
1	12. Is the tested exposure scenario relevant for the substance?	Fulfilled	Yes
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	Yes

Psammechinus miliaris larval development



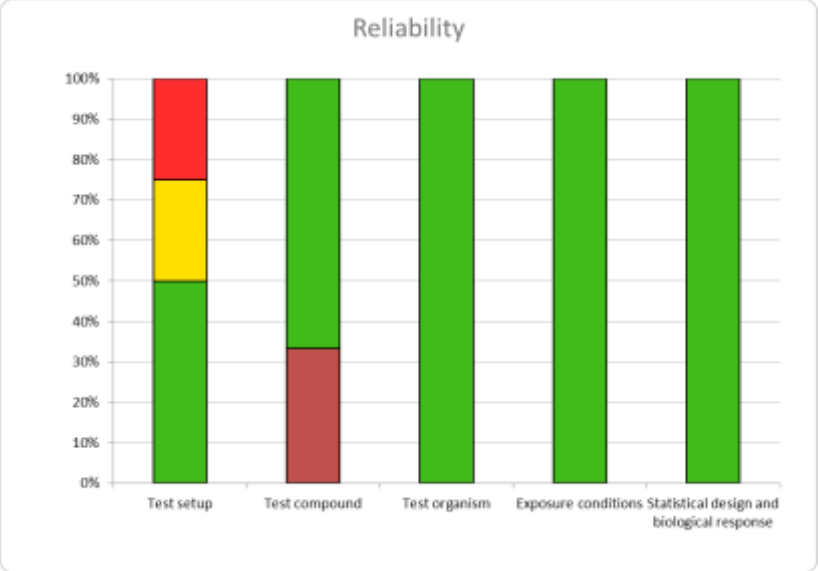
Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	3	15.79%	
Fulfilled	14	73.68%	
Partially fulfilled	2	10.53%	
Not fulfilled	0	0.00%	

Weight/Removed	Evaluation criteria	Selection	Comment
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Partially fulfilled	Fertlizations assay following the static toxicity test using sea urchin gametes adapted from EPS1 RM/27 (Fev.2011) Biological Test Method: Fertilization assay using echinoids (sea urchins and sand dollars). Larvae development following marine ecology and ecotoxicology reports/papers
1	2. Is the test performed under GLP conditions? (of minor importance forstudy reliability)	Partially fulfilled	Well described method for fertolization assay. Methods have been well described but not standardized for larvae development in *Psammechinus miliaris*
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	Control survival and normal larvae development
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	Positive and negative controls included
	Test compound		
1	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Not reported	Scrubber water is a mixture of pollutants
1	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Not reported	Scrubber water might be varied under different conditions
1	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not reported	Other compounds could be present and have an effect on organisms
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Yes. *Psammechinus miliaris*
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	Organisms were colletced in field and aclimatized for weeks in filtered seawater regularly tested for pollutants. Organisms kept at ambient/tested temperature and fed at libitum. No contaminants added before the experiment and pH of natural seawater.
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	Static conditions. Glass crystallization dishes (50 ml)were used as test vessels filled with 20 ml of experimental water. Glass material used in all the steps to avoid the union of the compounds to the walls of the containers.
1	11. Is the experimental system appropriate for the test organism (e.g., choiceof medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	Variables such as temperature, salinity, pH, and DO were measured at the beginning of the experiment.Tests were conducted for 72h at 15ºC and 16h:8h light: darkness photoperiod. Water levels were checked throughout the exposure period, and they were adjusted only when necessary. pH correctionwas made when necessary at the beginning of the experiment. No significant pH and OD variation should be observed during the exposure.
Removed	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	Yes, 10 dilutions: 0.001; 0.01; 0.1; 1.0; 2.0; 5.0; 10.0; 20.0; 40.0; 100.0%
1	14. Is the exposure duration defined?	Fulfilled	30min, 72h
1	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Fulfilled	Yes
Removed	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Not applicable	
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	Yes. 5 replicates per treatment. Each container contains enough embryos at the beginning to obtain at least 100 larvae at the end of the exposure (72h)
1	18. Are appropriate statistical methods used?	Fulfilled	LC50 using SPSS. One-way ANOVA and Student's t-test analyses were performed to compare exposure treatments and controls. LOECs and NOECs were determined using one side Dunnett's post hoc tests or the equivalent test for non-parametric analyses
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	Yes
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	Yes



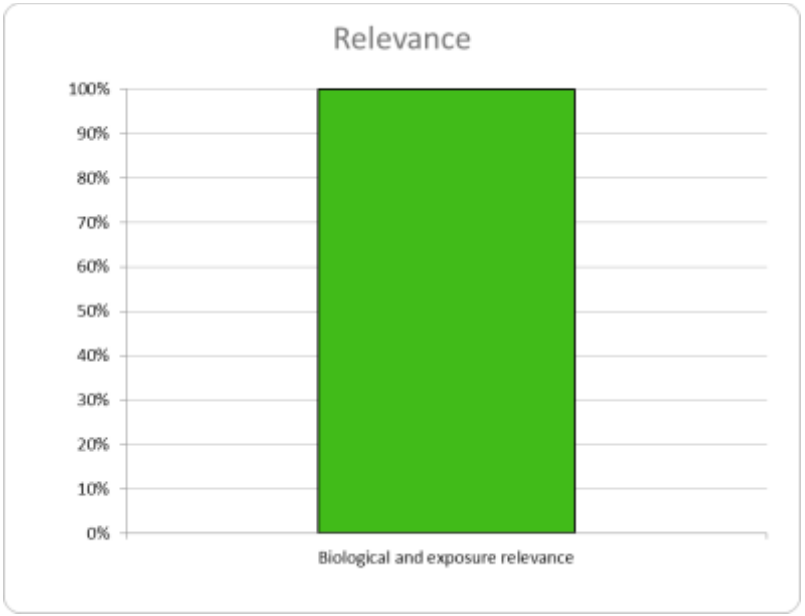
Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	0.00%	
Fulfilled	12	100.00%	
Partially fulfilled	0	0.00%	
Not fulfilled	0	0.00%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Yes, embryo-larval toxicity testing of marine pelagic larvae of *Psammechinus miliaris*
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, marine pelagic larvae of *Psammechinus miliaris*
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Fulfilled	Yes. Fertilization success, larval development
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	
1	6. Are appropriate life stages studied?	Fulfilled	Yes. Embryos and pelagic larvae
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	Yes
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	Yes, seawater
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	Yes, 30 min for fertilization and 72h for larvae
Removed	10. If recovery is studied, is this relevant for the framework for which the study is evaluated?	Not applicable	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	Yes, scrubber effluent
1	12. Is the tested exposure scenario relevant for the substance?	Fulfilled	Yes
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	Yes

Appendix 8: CRED analysis IVL



Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	1	5.26%	
Fulfilled	16	84.21%	
Partially fulfilled	1	5.26%	
Not fulfilled	1	5.26%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Not fulfilled	Marine ecology focus
1	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Partially fulfilled	Well described method but not standardized.
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	Control, survival, growth
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	No positive control, mainly relevant for standardised tests or single compound testing.
	Test compound		
1	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Fulfilled	Yes, scrubber effluent
1	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Fulfilled	Yes, Leo C scrubber effluent released under certain conditions
1	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not reported	The whole" formulation" or effluent is tested
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Yes. *Strongylocentrotus droebachiensis*
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	All is fulfilled. Acclimation for weeks in running seawater at ambient/tested temperature. No contaminants added before the experiment and pH of natural seawater.
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	Static conditions. test vessels were crystallization dishes (2-4 ml) filled with approximately 2 ml experimental water. 150 ml exposure water in glass flasks (red cap) for growth measurements.
1	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	Ecological aspects were targeted in these experiments. pH variations were part of the exposure.
Removed	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	Yes 10
1	14. Is the exposure duration defined?	Fulfilled	15 min, 11 d
1	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Fulfilled	Semi-static exposure for growth and static for fertilization.
1	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Fulfilled	Appropriate for the species and life stages tested.
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	Yes
1	18. Are appropriate statistical methods used?	Fulfilled	Anova, Permanova, Curve fit models, REGTOX

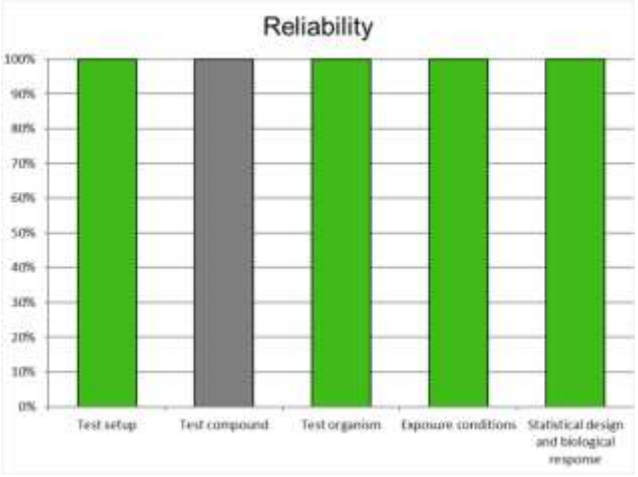
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	Yes
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	Yes



Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	0	0.00%	
Fulfilled	12	100.00%	
Partially fulfilled	0	0.00%	
Not fulfilled	0	0.00%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Yes, pelagic larvae
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, pelagic larvae
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Fulfilled	Yes, fertilization success, larval development, growth.
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	Yes. The tested endpoints are sensitive and are known to be affected by metals, PAHs and pH.
1	5. Is the effect relevant on a population level?	Fulfilled	Yes. It affects early life stages which are essential for the growth and sustainance of a population
1	6. Are appropriate life stages studied?	Fulfilled	Yes. Pelagic larvae
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	Yes
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	Yes. Seawater replaced intermittently to ensure good conditions.
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	Yes, 15 min for fertilization and 11 days with one day increments for growth and development
Removed	10. If recovery is studied, is this relevant for the framework for which the study is evaluated?	Not applicable	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	Yes, scrubber effluent
1	12. Is the tested exposure scenario relevant for the substance?	Fulfilled	Yes, especially for fertilization
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	Yes

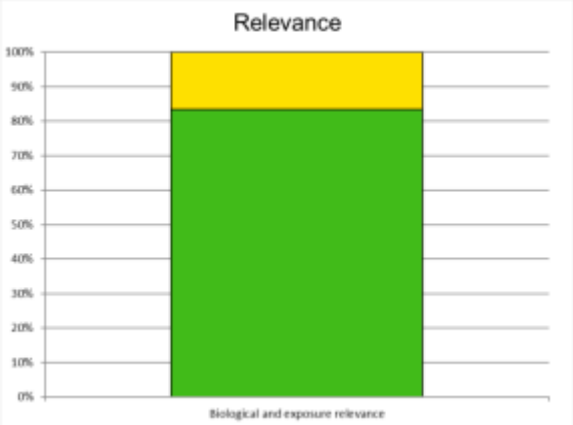
Appendix 9: CRED analysis UAV

Paracentrotus lividus 1



Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	2	11.11%	
Fulfilled	10	55.56%	
Partially fulfilled	5	27.78%	
Not fulfilled	1	5.56%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Fulfilled	Static toxicity test using sea urchin gametes and embryo. Adaptation from EPS1 RM/27 (Fev.2011) Biological Test Method: Fertilization assay using echinoids (sea urchins and sand dollars)
1	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Fulfilled	well described method
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	% of fertilization success
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	negative and positive control (SDS or CUSO4)
	Test compound		
1	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Not reported	No, the scrubber water is a mixture of contaminants
1	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Not reported	no
1	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not reported	The mixture exerts the effects
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Yes
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	Individuals of Paracentrotus lividus were sampled from a clean rocky beach a north of Aveiro, during the low tide and during natural spawning season. The animals were acclimated at 17 ± 1 ºC with a 18 h light:6 h darkness cycle in a recirculating natural seawater aquarium (salinity = 35 ± 1) for 3 d before starting the experiments. Sea urchins that are held in the laboratory for an extended period of time (i.e., > 3 days) should be fed with kelp or macroalga or with romaine lettuce (leafy greens and corn).
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	Yes
1	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	Tests were conducted for 20 minutes + 20 minutes. Water levels were checked throughout the exposure period, and they were adjusted only when necessary. Salinity, pH, temperature and DO were measured at the beginning and end of the test in the water samples and collected at the same time sub-samples for the metals and PAHs concentration analysis. pH correction and some times OD adjustment are necessary at the beginning. No significant pH and OD variation should be observed during the exposure.
Removed	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	appropriate according to a factor of: 2 or 10^x
1	14. Is the exposure duration defined?	Fulfilled	20min+20min

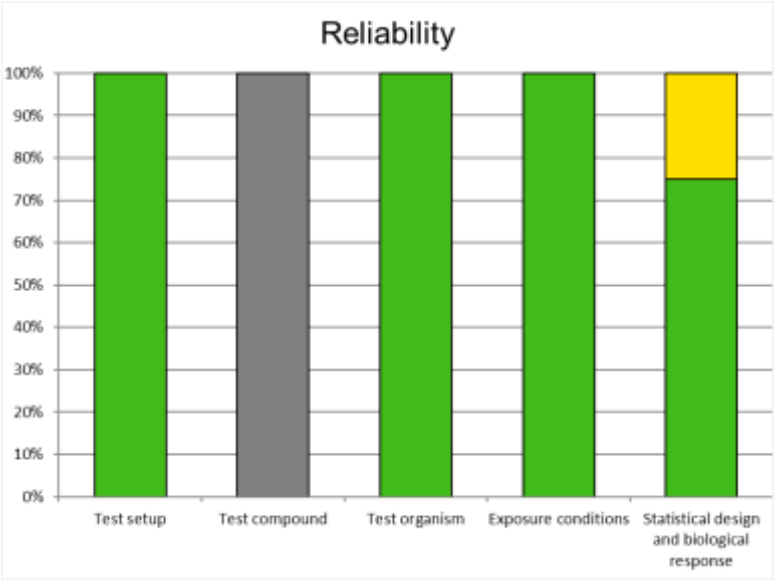
1	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Fulfilled	Yes
Removed	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Not applicable	
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	Each trial was composed of a minimum of 5 exposure concentrations plus the control and carried out in 5 replicates. Sperm that fertilized 95% of the oocytes. 20-30 oocites/mL were used per vessel of 10mL
1	18. Are appropriate statistical methods used?	Fulfilled	
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Fulfilled	Yes
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	



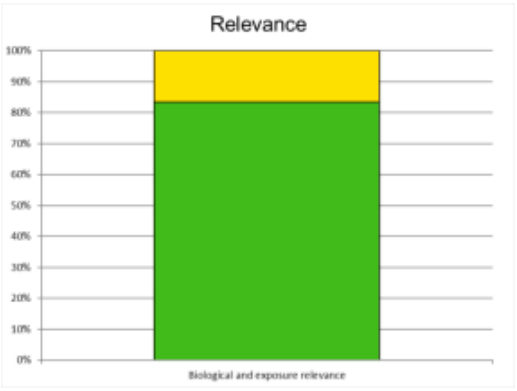
Evaluation result	Total, accounting for weight		%, accounting for weight
Not determined		0	0.00%
Not reported		0	0.00%
Fulfilled		9	75.00%
Partially fulfilled		3	25.00%
Not fulfilled		0	0.00%

Weight/Removed	Evaluation criteria	Selection	Comment
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Paracentrotus lividus - gametes>embryo
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, the developmental stage is pelagic (gametes are released into the seawater)
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Partially fulfilled	fertilization success (appropriate to the regulatory ??)
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	
1	6. Are appropriate life stages studied?	Fulfilled	
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	20min+20min
Removed	10. If recovery is studied, is the relevant for the therapeutic for which the study is evaluated?	Not assessed	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	?
1	12. Is the tested exposure scenario relevant for the substance?	Partially fulfilled	?
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	Yes, the developmental stage is pelagic (gametes are released into the seawater)

Paracentrotus lividus 2



Evaluation result	Total, accounting for weight	%, accounting for weight	
Not determined	0	0.00%	
Not reported	2	11.11%	
Fulfilled	10	55.56%	
Partially fulfilled	5	27.78%	
Not fulfilled	1	5.56%	
Weight/Removed	**Evaluation criteria**	**Selection**	**Comment**
	Test setup		
1	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used? (of minor importance for study reliability)	Fulfilled	Static toxicity test using sea urchin embryo and larvae. Adaptation from EPS1 RM/58 (Jul.2014) Reference method for measuring the toxicity of contaminated sediment to embryo and larvae of echinoids (sea urchin and sand dollars) and ISSO 17244: 2015- Determination of the toxicity of water samples on the embryo-larval development of Japanese oyster (Crassostrea gigas) and mussel (Mytilus galloprovincialis)
1	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Fulfilled	well described method
1	3. If applicable, are validity criteria fulfilled (e.g. control survival, growth)?	Fulfilled	control survival and % of abnormality
1	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	negative and positive control
	Test compound		
1	5. Is the test substance identified clearly with name or CAS-number? Are test results reported for the appropriate compound?	Not reported	No, the scrubber water is a mixture of contaminants
1	6. Is the purity of the test substance reported? Or, is the source of the test substance trustworthy?	Not reported	no
1	7. If a formulation is used or if impurities are present: Do other ingredients in the formulation exert an effect? Is the amount of test substance in the formulation known?	Not reported	The mixture exerts the effects
	Test organism		
1	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, strain/clone, gender if appropriate)?	Fulfilled	Yes
1	9. Are the test organisms from a trustworthy source and acclimatized to test conditions? Have the organisms not been pre-exposed to test compound or other unintended stressors?	Fulfilled	Individuals of Paracentrotus lividus were sampled from a clean rocky beach a north of Aveiro, during the low tide and during natural spawning season. The animals were acclimated at 17 ± 1 ºC with a 18 h light:6 h darkness cycle in a recirculating natural seawater aquarium (salinity = 35 ± 1) for 3 d before starting the experiments. Sea urchins that are held in the laboratory for an extended period of time (i.e., > 3 days) should be fed with kelp or macroalga or with romaine lettuce (leafy greens and corn).
	Exposure conditions		
1	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical characteristics?	Fulfilled	Yes
1	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or test water, feeding, water characteristics, temperature, light/dark conditions, pH, oxygen content)? Have conditions been stable during the test?	Fulfilled	Tests were conducted for 48h at 17-20ºC and a 16 h light:08 h darkness photoperiod. Water levels were checked throughout the exposure period, and they were adjusted only when necessary. Salinity, pH, temperature and DO were measured at the beginning and end of the test in the water samples and collected at the same time sub-samples for the metals and PAHs concentration analysis. pH correction and some times OD adjustment are necessary at the beginning. No significant pH and OD variation should be observed during the exposure.
Removed	12. Were exposure concentrations below the limit of water solubility (taking the use of a solvent into account)? If a solvent is used, is the solvent within the appropriate range and is a solvent control included?	Not applicable	
1	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	appropriate according to a factor of 2 or 10^x
1	14. Is the exposure duration defined?	Fulfilled	48h
1	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Fulfilled	Yes
Removed	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Not applicable	
	Statistical design and biological response		
1	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls and test concentrations?	Fulfilled	Each trial was composed of a minimum of 5 exposure concentrations plus the control and carried out in 5 replicates. 20-30 embryos/mL were used per vessel of 20mL
1	18. Are appropriate statistical methods used?	Fulfilled	Yes
1	19. Is a concentration-response curve observed? Is the response statistically significant?	Partially fulfilled	Yes (in our experience we need to expand the range to lower concentrations).
1	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control data, concentration-response curves)?	Fulfilled	

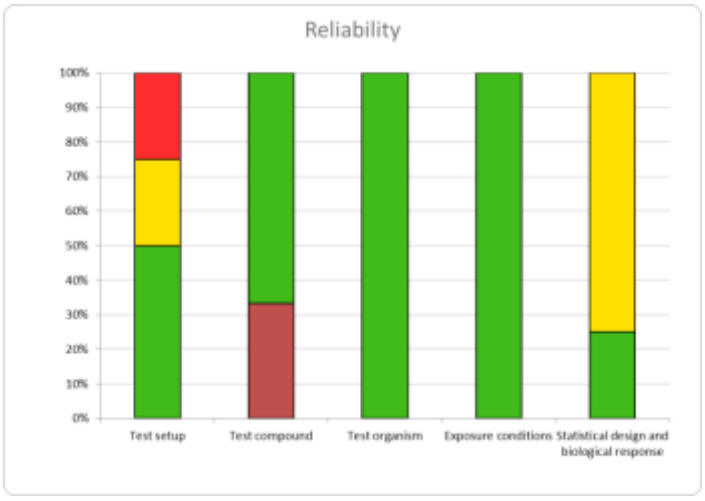


Evaluation result	Total, accounting for weight	% accounting for weight
Not determined	0	0.00%
Not reported	0	0.00%
Fulfilled	3	75.00%
Partially fulfilled	1	25.00%
Not fulfilled	0	0.00%

Weight/Removed	Evaluation criteria	Selection	Comment
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Paracentrotus lividus - embryo and larvae stages very sensible
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes,the development stage is pelagic (live in the water)
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Partially fulfilled	mortality (partly appropriate) and abnormality responses (appropriate)
1	4. Are the reported endpoints appropriate for the investigated effects or the mode of action of the test substance?	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	
1	6. Are appropriate life stages studied?	Fulfilled	
1	7. Is the magnitude of effect statistically significant and biologically relevant for the regulatory purpose (e.g., EC10, EC50)?	Fulfilled	
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	
1	9. Is the exposure duration relevant and appropriate for the studied endpoints and species?	Fulfilled	48h
[illegible]	[illegible]	[illegible]	
1	11. In case of a formulation, other mixture, salts, or transformation products, is the substance tested representative and relevant for the substance being assessed?	Fulfilled	
1	12. Is the tested exposure scenario relevant for the substance?	Partially fulfilled	)
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	Yes, the developmental stage is pelagic (gametes are released into the seawater)

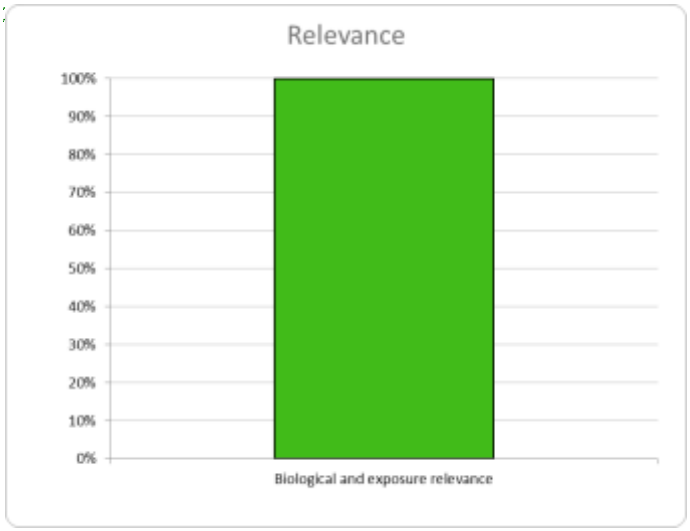
Appendix 10: CRED analysis AUTH

Phytoplankton & bacterioplankton



Evaluation result:	Total, accounting for weight	% accounting for weight
Not determined		0 0.00%
Not reported		1 5.26%
Fulfilled		12 63.16%
Partially fulfilled		4 21.05%
Not fulfilled		1 5.26%

Weight/Removed	Evaluation criteria	Selection	Comment
	Test setup		
5	1. Is a guideline method (e.g., OECD/ISO) or modified guideline used (or minor importance for study reliability)	Not fulfilled	marine ecology focus
5	2. Is the test performed under GLP conditions? (of minor importance for study reliability)	Partially fulfilled	well-described method but not standardized
5	3. If applicable, are validity criteria fulfilled (e.g., control survival, growth)?	Fulfilled	(control, survival, growth)
5	4. Are appropriate controls performed (e.g. solvent control, negative and positive control)?	Fulfilled	Negative control and positive Nitrate nitrogen control performed. For the Thermikos Gulf experiment
	Test compound		
5	5. Is the test substance identified clearly with name or CAS number? Are test results reported for the appropriate	Fulfilled	Yes, a rubber effluent
5	6. Is the purity of the test substance reported? Or, if the source of the test substance is unknown?	Fulfilled	Yes, I-or-C a rubber effluent released under certain conditions
5	7. If a formulation is used or if impurities are present. Do other ingredients in the formulation exert an effect? Is the	Not reported	The whole "formulation" or effluent is tested
	Test organism		
5	8. Are the organisms well described (e.g. scientific name, weight, length, growth, age/life stage, sex/culture, gender if	Fulfilled	Yes. Phytoplankton and bacterioplankton several species
5	9. Are the test organisms from a trustworthy source and acclimated to test conditions? Have the organisms not	Fulfilled	All is fulfilled. Seawater at ambient/tested temperature. No contaminants added before the experiment and pH of natural seawater.
	Exposure conditions		
5	10. Is the experimental system appropriate for the test substance, taking into account its physico-chemical	Fulfilled	Static conditions with gentle mixing. Test vessels were glass bottles (1L) filled with .1L. experimental seawater.
5	11. Is the experimental system appropriate for the test organism (e.g., choice of medium or soil-water, feeding,	Fulfilled	Ecological aspects were subjected in these experiments, yet variations in short-time natural range were part of the exposure.
	[illegible]		
5	13. Is a correct spacing between exposure concentrations applied?	Fulfilled	Yes, between 2 and 30
5	14. Is the exposure duration defined?	Fulfilled	1 and 7 days in combination with 24 hours examinations
5	15. Are chemical analyses adequate to verify concentrations of the test substance over the duration of the study?	Fulfilled	72 hrs. exposure for growth with gentle stir 'n' dry mixing
5	16. Is the biomass loading of the organisms in the test system within the appropriate range (e.g. < 1 g/L)?	Fulfilled	Appropriate for the species tested
	Statistical design and biological response		
5	17. Is a sufficient number of replicates used? Is a sufficient number of organisms per replicate used for all controls	Fulfilled	Yes
5	18. Are appropriate statistical methods used?	Partially fulfilled	Mann-Whitney non-parametric test
5	19. Is a concentration-response curve observed? Is the response statistically significant?	Partially fulfilled	Yes, the response was significant
5	20. Are sufficient data available to check the calculation of endpoints and (if applicable) validity criteria (e.g., control	Partially fulfilled	Three replicates



Evaluation result	Total, accounting for weight	% accounting for weight
Not determined	0	0.00%
Not reported	0	0.00%
Fulfilled	11	100.00%
Partially fulfilled	0	0.00%
Not fulfilled	0	0.00%

Weight/Removed	Evaluation criteria	Selection	Comment
	Biological and exposure relevance		
1	1. Is the species tested relevant for the compartment under evaluation?	Fulfilled	Yes, pelagic algae and bacteria
1	2. Are the organisms tested relevant for the tested substance?	Fulfilled	Yes, pelagic algae and bacteria
1	3. Are the reported endpoints appropriate for the regulatory purpose?	Fulfilled	Yes, growth rate, numbers
1	4. Are the reported endpoints appropriate for the investigated effects or the	Fulfilled	
1	5. Is the effect relevant on a population level?	Fulfilled	
Removed	6. Are appropriate life stages studied?	Not applicable	
1	7. Is the magnitude of effect statistically significant and biologically relevant for	Fulfilled	Yes
1	8. Are the experimental conditions relevant for the tested species?	Fulfilled	Yes, seawater
1	9. Is the exposure duration relevant and appropriate for the studied endpoints	Fulfilled	Yes, 24 hours for acute and 3 days for acute/chronic growth
Removed	10. If recovery is studied, is the relevant for the framework for which the study	Not applicable	
1	11. In case of a formulation, other mixture, salts, or transformation products, is	Fulfilled	Yes, scrubber effluent
1	12. Is the tested exposure scenario relevant for the substance?	Fulfilled	Yes
1	13. Is the tested exposure scenario relevant for the species?	Fulfilled	Yes

Appendix 11: COMPILATION OF ECOTOX DATA, EMERGE, 2022-10-25

Institute	Scrubber water origin	Area	Test category	Test description	Species	Species	Life stage	Endpoint	Lowest test [C] (% dilution)	LOEC (% dilution)	NOEC (% dilution)	EC10 (% dilution)	LC50 (% dilution)
UV	Chalmers	Mediterranean	acute	bioluminescence	bacteria	Aliivibrio fisheri	bacteria	bioluminescence	0.01	20	10	24.8	
UV	Chalmers	Mediterranean	acute	Acute 48 h lethality	copepod	Acartia tonsa	adult	mortality	0.01	40	20	36.4	
UV	Chalmers	Mediterranean	chronic	Life-cycle test	copepod	Acartia tonsa	egg	hatching success	0.01	40	20	n.c.	
UV	Chalmers	Mediterranean	chronic	Life-cycle test	copepod	Acartia tonsa	Nauplia VI	survival	0.01	20	10	9.1	
UV	Chalmers	Mediterranean	chronic	LDR 5 day test	copepod	Acartia tonsa	egg to copepodite	larval development	0.01	0.01	<0.01	n.c.	
UV	Chalmers	Mediterranean	chronic	Long term exposure	copepod	Acartia tonsa	adult with eggs	egg production	0.01	0.01	<0.01	n.c.	
UV	DANAOS-Leo C_10B	Mediterranean	acute	bioluminescence	bacteria	Aliivibrio fisheri	bacteria	bioluminescence	0.01	20	10	23	39
UV	DANAOS-Leo C_10B	Mediterranean	acute	acute 48 h lethality	copepod	Acartia tonsa	adult	mortality	0.01	10	5	8	11
UV	DANAOS-Leo C_10B	Mediterranean	acute	algal growth test	microalga	Phaeodactylum tricornutum	microalgae	growth rate	5	40	20	34	n.c.
UV	DANAOS-Leo C_10B	Mediterranean	acute	algal growth test	microalga	Dunaliella tertiolecta	microalgae	growth rate	5	20	10	15	n.c.
UV	DANAOS-Leo C_10B	Mediterranean	chronic	larval development bioassay	mussel	Mytilus galloprovincialis	embryos	larval development	0.001	1	0.1	4.9	6
UV	DANAOS-Leo C_10B	Mediterranean	chronic	Life-cycle test	copepod	Acartia tonsa	adult (F0)	egg production	0.001	0.01	0.001	n.c.	n.c.
UV	DANAOS-Leo C_10B	Mediterranean	chronic	Life-cycle test	copepod	Acartia tonsa	eggs (F1)	hatching success	0.001	> 1	> 1	n.c.	n.c.
UV	DANAOS-Leo C_10B	Mediterranean	chronic	Life-cycle test	copepod	Acartia tonsa	larvae (F1)	larval survival	0.001	> 1	> 1	n.c.	n.c.
UV	DANAOS-Leo C_10B	Mediterranean	chronic	Life-cycle test	copepod	Acartia tonsa	larvae (F1)	larval development	0.001	0.1	0.01	n.c.	n.c.
UV	DANAOS-Leo C_10B	Mediterranean	chronic	LDR 5 day test	copepod	Acartia tonsa	from egg to copepodite	hatching success	0.01	20	10	9	25
UV	DANAOS-Leo C_10B	Mediterranean	chronic	LDR 5 day test	copepod	Acartia tonsa	from egg to copepodite	larval survival	0.01	20	10	9	13
UV	DANAOS-Leo C_10B	Mediterranean	chronic	LDR 5 day test	copepod	Acartia tonsa	from egg to copepodite	larval development	0.01	2	1	1.1	1.5
UoS	Chalmers	Mediterranean	chronic	Fertilization success	mussel	Mytilus edulis	embryo	fertilization success	0.001	1	0.1	0.58	40.89
UoS	Chalmers	Mediterranean	chronic	Larval development bioassay	mussel	Mytilus edulis	larvae	abnormal larval development	0.001	0.001	<0,001	0.06	0.54
UoS	Chalmers	Mediterranean	chronic	Fertilization success	sea urchin	Psammechinus miliaris	embryo	fertilization success	0.001	0.01	0.001	0.12	1.2
UoS	Chalmers	Mediterranean	chronic	Larval development bioassay	sea urchin	Psammechinus miliaris	larvae	abnormal larval development	0.001	0.01	0.001	0.11	0.96
UoS	DANAOS-Catherine C	North Sea	chronic	Fertilization success	mussel	Mytilus edulis	embryo	fertilization success	0.001	2	1	0.49	156
UoS	DANAOS-Catherine C	North Sea	chronic	Larval development bioassay	mussel	Mytilus edulis	larvae	abnormal larval development	0.001	0.001	<0,001	0.27	9.27
UoS	DANAOS-Catherine C	North Sea	chronic	Fertilization success	sea urchin	Psammechinus miliaris	embryo	fertilization success	0.001	1	0.1	0.12	4.31
UoS	DANAOS-Catherine C	North Sea	chronic	Larval development bioassay	sea urchin	Psammechinus miliaris	larvae	abnormal larval development	0.001	0.1	0.01	0.15	1.54
IVL	DANAOS-Leo C_1B	North Atlantic	chronic	Larval development	sea urchin	Strongylocentrotus droebachiensis	larvae	abnormal larval count	0.0001	0.1	0.01	2.677	4.678
IVL	DANAOS-Leo C_1B	North Atlantic	acute	Fertilization success	sea urchin	Strongylocentrotus droebachiensis	embryo	fertilization success	0.0001	0.0001	<0.0001	2.339	7.708
IVL	Thor et al. 2021	Skagerrak/Oresund	chronic	Chronic	copepod	Calanus helgolandicus	copepodite (CIII)	moulting	1	1	<1		
IVL	Thor et al. 2021	Skagerrak/Oresund	chronic	Chronic	copepod	Calanus helgolandicus	copepodite (CIII)	mortality	1	1	<1		
UAV	Chalmers	Atlantic	chronic	Fertilization success	sea urchin	Paracentrotus lividus	embryo	fertilization success	0.01	0.01	<0,01	6.36	26.68
UAV	Chalmers	Atlantic	chronic	Larval development bioassay	sea urchin	Paracentrotus lividus	larvae	abnormal larval development	0.001	0.01	0.001	0.265	8.04
UAV	Chalmers	Atlantic	chronic	Larval development bioassay	polychaete	Sabellaria alveolata	larvae	abnormal larval development	0.001	0.01	0.001	<0,001	3.8
UAV	Chalmers	Atlantic	chronic	Fertilization success	sea urchin	Paracentrotus lividus	embryo	fertilization success	1.56	3.125	1.56		13.7
UAV	Chalmers	Atlantic	chronic	Larval development bioassay	sea urchin	Paracentrotus lividus	larvae	abnormal larval development	1.56	1.56	<1.56		1.3
UAV	DANAOS-Catherine C	Atlantic	chronic	Fertilization success	sea urchin	Paracentrotus lividus	embryo	fertilization success	1.56	3.125	1.56		22.9
UAV	DANAOS-Catherine C	Atlantic	chronic	Larval development bioassay	sea urchin	Paracentrotus lividus	larvae	abnormal larval development	1.56	1.56	<1.56		1.5
UAV	DANAOS-Catherine C	Atlantic	chronic	Fertilization success	sea urchin	Paracentrotus lividus	embryo	fertilization success	0.01	0.01	<0,01	7.56	33.66
UAV	DANAOS-Catherine C	Atlantic	chronic	Larval development bioassay	sea urchin	Paracentrotus lividus	larvae	abnormal larval development	0.001	0.001	<0,001	-	6.13
UAV	DANAOS-Catherine C	Atlantic	chronic	Larval development bioassay	polychaete	Sabellaria alveolata	larvae	abnormal larval development	0.001	0.01	0.001	-	9.44
UAV	DANAOS-Leo C_1B	North Atlantic	chronic	Fertilization success	sea urchin	Paracentrotus lividus	embryo	fertilization success	0.01	0.1	0.01	7.22	11.38
UAV	DANAOS-Leo C_1B	North Atlantic	chronic	Larval development bioassay	sea urchin	Paracentrotus lividus	larvae	abnormal larval development	0.001	0.01	0.1	0.78	5.51
UAV	DANAOS-Leo C_1B	North Atlantic	chronic	Larval development bioassay	polychaete	Sabellaria alveolata	larvae	abnormal larval development	0.001	0.001	<0,001	1.13	10.47
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Pseudonitzschia cf. pungens	-	population density	1	10	<10, >1	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Heterocapsa rotundata	-	population density	1	10	<10, >1	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Chrysochromulina sp.	-	population density	1	10	<10, >1	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Teleaulax sp.	-	population density	1	10	<10, >1	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute	Acute 24h	-	Phytoplankton community	-	total abundance	1	-	-	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	chronic	Chronic 72 h	-	Phytoplankton community	-	total abundance	1	-	-	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Cylindrotheca closterium	-	population density	1	-	>10	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Gymnodinium sp.	-	population density	1	10	<10, >1	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Chrysochromulina sp.	-	population density	1	10	<10, >1	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Teleaulax sp.	-	population density	1	10	<10, >1	-	-

AUTH	DANAOS-Leo C_11B	Mediterranean	acute	Acute 24h	-	Phytoplankton community	-	total abundance	1	-	-	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	chronic	Chronic 72 h	-	Phytoplankton community	-	total abundance	1	-	-	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Skeletonema sp.	-	population density	1	5	<5, >2	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Gymnodinium sp.	-	population density	1	5	<5, >2	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Chrysochromulina sp.	-	population density	1	5	<5, >2	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Teleaulax sp.	-	population density	1	2	<2, >1	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute	Acute 24h	-	Phytoplankton community	-	total abundance	1	-	-	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	chronic	Chronic 72 h	-	Phytoplankton community	-	total abundance	1	-	-	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Skeletonema sp.	-	population density	1	-	>5	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Gymnodinium sp.	-	population density	1	5	<5, >2	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Chrysochromulina sp.	-	population density	1	5	<5, >2	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute/chronic	Acute/Chronic 72 h	microalga	Teleaulax sp.	-	population density	1	5	<5, >2	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	acute	Acute 24h	-	Phytoplankton community	-	total abundance	1	-	-	-	-
AUTH	DANAOS-Leo C_11B	Mediterranean	chronic	Chronic 72 h	-	Phytoplankton community	-	total abundance	1	-	-	-	-