SOM topic report: Hazardous substances

HELCOM ACTION project, HELCOM Secretariat

28 October 2020

Table of Contents

Summary of main results of the analysis

Existing measures are not sufficient in achieving GES for mercury, TBT, PFOS and diclofenac.

Probability to achieve GES with existing measures is very low for mercury, TBT and diclofenac, and low for PFOS. The overall certainty of this assessment is moderate.

Total pressure reductions from existing measures are zero for TBT and diclofenac, and from low to moderate for mercury and PFOS.

There is considerable uncertainty about the total pressure reductions required to achieve GES.

Main pressures contributing to the concentration of the substances are:

> Concentration of mercury: Heavy metal pollution; Change in hydrologic conditions
>
> Concentration of PFOS: Organohalogen pollution (e.g. PFOS, PCBs, PBDEs, dioxins); River, lake, or land habitat loss/degradation
>
> Concentration of TBT: Organotin pollution (e.g. TBT); Physical disturbance of marine habitats
>
> Concentration of diclofenac: Pharmaceutical pollution; Change in hydrologic conditions

Measure types having the most impact on the input of the substance are:

> Input of mercury: Local/state/national targets to eliminate coal fired energy production; Minamata convention; Paris agreement
>
> Input of PFOS: Clean-up of contaminated sites; Stockholm convention PFOS lists no accepted uses or exemptions; Restoration/upgrading of old landfill sites
>
> Input of TBT: Reduce re-suspension from sediments, by limiting restructuring of seabed to areas with low concentrations; Tighter allowed contamination levels during dredging activities; Implement technologies for the degradation of TBT from sediment (e.g. chemical oxidation, bioremediation)
>
> Input of diclofenac: Technical upgrade of wastewater treatment plants: e.g. granular activated carbon (GAC) adsorption, ozonation, UV light, nanofiltration etc.; Alter prescription practices to lower consumption (drug dosage, pack size, alternative medicine, convert OTC access to prescription)

Main activities contributing to the input of the substance are:

> Input of mercury: activities and sources outside the Baltic Sea Region; non-renewable energy generation; industrial uses
>
> Input of PFOS: solid waste; waste waters; activities and sources outside the Baltic Sea Region
>
> Input of TBT: solid waste; restructuring of seabed morphology
>
> Input of diclofenac: waste waters; solid waste

Background

Report background

The sufficiency of measures (SOM) analysis assesses improvements in environmental state and pressures that can be achieved with existing measures in the Baltic Sea, and whether these are sufficient to achieve good environmental status (GES). The analysis involves estimating the state of the marine environment in 2030, given measures in existing policies, their implementation status and projected development of human activities over time, which can be compared to the agreed HELCOM threshold for GES, when available.

The main aim of the SOM analysis is to support the update of the HELCOM Baltic Sea Action Plan (BSAP) by identifying potential gaps in achieving environmental objectives with existing measures for the Baltic Sea. In addition, the analysis can indicate both thematically and spatially where new measures are likely needed.

The same overall approach has been applied across all topics included in the SOM analysis to ensure comparability and coherence of the results, while considering topic-specific aspects and making necessary adjustments. The main components of the analysis include assessing the contribution of activities to pressure inputs, the effect of existing measures on pressure inputs, the effect of development of human activities on pressure inputs, and the effect of changes in pressure on environmental state. The SOM approach, model and data collection are described in detail in the methodology report.

The methodology for the SOM analysis is designed to accommodate for the broad array of topics relevant in the HELCOM region and to enable a region-level analysis. It balances between state-of-the-art knowledge, availability of data, and advice taken onboard from various HELCOM meetings and bodies.

The data used in the SOM analysis have been collected using expert elicitation and by reviewing existing literature, model outputs and other data sources. Data availability varies substantially across topics and data components, which is reflected in the presentation of the methods and results in this report.

The SOM analysis presents the first attempt to quantify the effects of existing measures and policies on the environment and achieving policy objectives for various environmental topics in HELCOM and the Baltic Sea area. It is aimed at assessing the overall sufficiency of existing measures at the Baltic Sea level. The results are based mainly on expert elicitation, and thus they should be considered as approximate. Due to the pioneering nature and variable data quality and availability of the SOM analysis, the findings do not provide complete or final answers on the need for new measures, and should be reviewed in relation to the results of other assessments.

This topic report describes the analyses and results for hazardous substances in the SOM analysis, providing detailed topic-specific information. First, it presents background information and describes the data and methods for addressing the topic in the SOM assessment, including relevant assumptions and challenges. Second, it presents and discusses the findings for each result component. Third, it provides discussion on the impacts of alternative assumptions and data, evaluates the quality and confidence of the analysis, and provides implications and future perspectives. The annexes contain detailed information on the data components, topic structure and expert surveys for the analysis, as well as supplementary results.

Similar topic reports will be prepared for all nine topics covered in the SOM analysis. In addition, the results are summarized in the main report and the full methodology is described in the methodology report.

Topic background

The Baltic Sea is heavily polluted by hazardous substances originating from human activities (HELCOM, 2018f). Exposure of living organisms to excessive levels of toxic substances can lead to difficulties to function, grow and reproduce, or even death. This can in turn lead to loss of biodiversity, ecosystem functionality, and ecosystem services. Toxic substances that degrade slowly and accumulate within the food web are especially problematic since they remain in the environment for many years. Some compounds are well-known, but many are largely unstudied and only a small subset of the immense number produced are monitored (Sobek et al., 2016). Hazardous substances known to pollute the Baltic Sea belong to chemical classes or use categories such as organotins, organohalogens, heavy metals, pharmaceuticals, polycyclic aromatic hydrocarbons, and radioactive substances.

Organotins are tin-based compounds with hydrocarbon substituents, which have historically been used e.g. in antifouling paints and wood preservatives (Larsen, 2020). Tin is not particularly toxic, but with the organic functional groups, the toxicity can become very high. The well-known compound TBT is a known hormone disruptor causing e.g. sex abnormalities in oysters (imposex). The degradation rate of TBT is low, especially under anoxic conditions. Today, the use of TBT in antifouling paints and wood preservatives is forbidden in the EU, and there are ongoing efforts to phase out other organotins as well. TBT may however still be released from older paint layers on ship hulls, as well as, from contaminated sediments during activities such as dredging.

Organohalogens are organic compounds that contain at least one halogen (e.g fluorine, chlorine, bromine) bonded to carbon. One example is per- and polyfluorinated alkyl substances (PFAS), which are commonly used in textiles, fire-fighting foam, hydraulic oil, and ski waxes (Johansson and Undeman, 2020). Since many PFAS are very persistent, bioaccumulative and toxic (Blom and Hansen, 2015), there are ongoing efforts to phase out the production and use, as outlined by the Stockholm Convention (Johansson and Undeman, 2020). The use of perfluorooctane sulfonate (PFOS), the PFAS that historically has been produced in the highest quantity, is specifically restricted. However, the reduction of PFOS use has led to increased use of other PFAS, which seem to have similar properties.

Heavy metals occur naturally in all environmental compartments, but concentrations have increased due to human activity (HELCOM, 2018g). The heavy metal assessed in this analysis is mercury, a toxic, persistent pollutant that in its methylated form biomagnifies through food webs (WHO, 2008). The main inputs of mercury to the Baltic Sea is through atmospheric deposition and via rivers (HELCOM, 2018g). Over the years, there have been many efforts to minimize mercury use within HELCOM and EU and further measures are outlined in the Minamata convention, which is implemented 2020 (Larsen, 2020).

Pharmaceuticals are useful chemicals for treatment of diseases but can also cause harm once released in the environment. They may enter the Baltic Sea mainly due to poor removal in wastewater treatment plants or due to improper disposal of unused medicines down the drain or veterinary use (UNESCO & HELCOM 2017). One of the most monitored and detected pharmaceuticals worldwide is the anti-inflammatory drug diclofenac, which has also been detected at several locations in the Baltic Sea (Undeman, 2020). Oil spills, combustion of fossil fuels, and waste incineration are sources of polycyclic aromatic hydrocarbons (PAHs) in the Baltic Sea (HELCOM, 2018h). These compounds are toxic and accumulate in aquatic organisms, particularly invertebrates. PAHs associate with particles in the water and accumulate in sediments where they can persist for a long time, especially in anaerobic sediments (HELCOM, 2018h).

After the accident at the Chernobyl nuclear power plant in 1986, the radionuclide cesium-137 (137Cs) was deposited over the Baltic Sea and has accumulated in marine biota and sediments (HELCOM, 2018i). The level of 137Cs is still higher than in any other ocean but is expected to decline below the threshold for good status in the Baltic Sea (pre-Chernobyl levels) in the coming few years since new inputs are low (HELCOM, 2018i).

Description of hazardous substances in the SOM assessment

Hazardous substances have been considered in a variety of ways in the SOM analysis. The primary focus is on the full analysis of four substances which are evaluated at each step of the SOM analysis. These substances are mercury, perfluorooctane sulphonate (PFOS), tributyltin (TBT), and diclofenac, which were selected to represent a variety of chemical attributes and management regimes. The input of these substances is tracked by the pressure inputs: *Input of mercury, Input of PFOS, Input of TBT, Input of diclofenac* (Figure 1), input being defined as new emissions to the environment or release/resuspension of loads otherwise removed from the system (e.g. dredging releasing buried TBT) caused by human activity. The concentrations of these substances are then tracked as the state components: *Concentration of mercury, Concentration of PFOS, Concentration of TBT, Concentration of diclofenac* (Figure 1). These state components reflect the structure of the HELCOM indicators "Metals (lead, cadmium and mercury)", "Perfluorooctane sulphonate", "TBT and imposex", and "Diclofenac", as well as MSFD criteria D8C1[1]. These state components have established HELCOM GES thresholds (mercury, PFOS) or test thresholds (TBT, diclofenac) and are evaluated only for their primary testing matrices[2] (water, sediment, biota) (HELCOM 2018b-e).

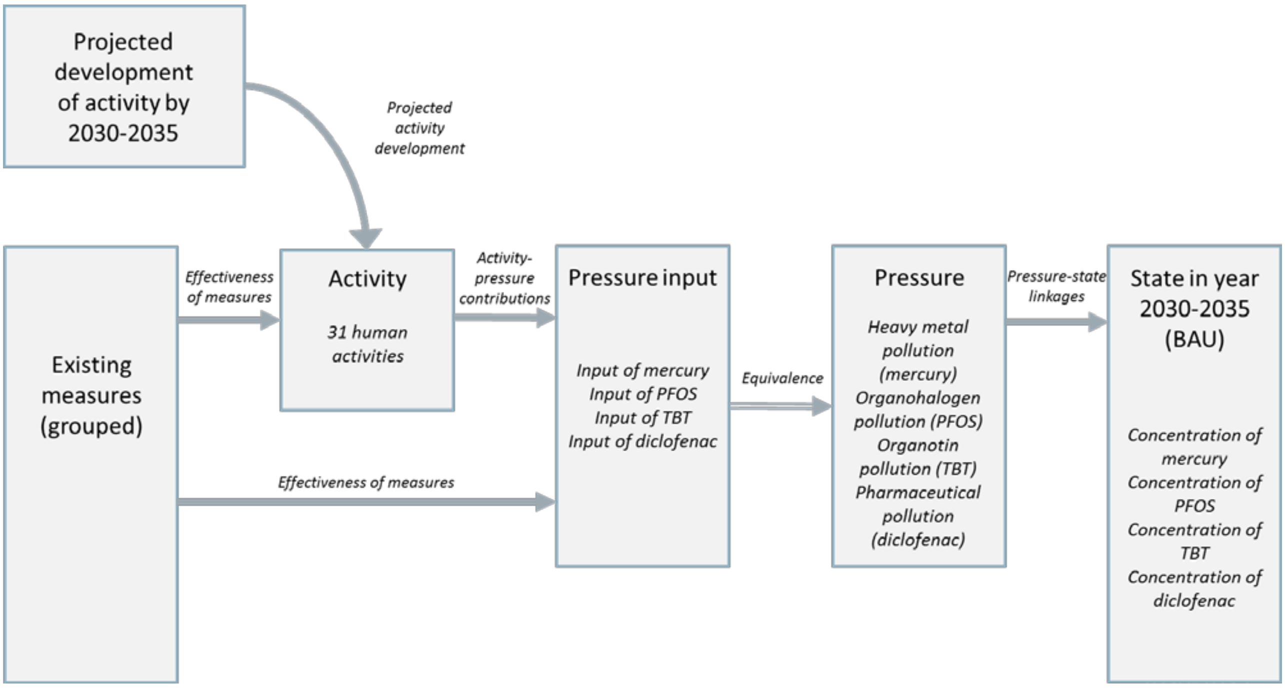

Figure 1. Schematic of the SOM analysis for the four analysed hazardous substances (mercury, PFOS, TBT, diclofenac). The pressure input (e.g. input of mercury) and the pressure (e.g. mercury pollution) relevant to each substance are assumed to be equivalent.

For mercury, the primary test matrix is fish muscle where the GES threshold is set at 20 µg/kg wet weight. In the latest HOLAS assessment period (2011-2016), mercury was in a not good status at the Baltic Sea scale; with good status achieved only in a few smaller scale areas in the southwestern Baltic (HELCOM 2018a-b).

[1] Marine Strategy Framework Directive criteria D8C1 – Primary: Within coastal and territorial waters, the concentrations of contaminants do not exceed the following threshold values:
(a) for contaminants set out under point 1(a) of criteria elements, the values set in accordance with Directive 2000/60/EC;
(b) when contaminants under point (a) are measured in a matrix for which no value is set under Directive 2000/60/EC, the concentration of those contaminants in that matrix established by Member States through regional or subregional cooperation;
(c) for additional contaminants selected under point 1(b) of criteria elements, the concentrations for a specified matrix (water, sediment or biota) which may give rise to pollution effects. Member States shall establish these concentrations through regional or subregional cooperation, considering their application within and beyond coastal and territorial waters.
Beyond territorial waters, the concentrations of contaminants do not exceed the following threshold values:
(a) for contaminants selected under point 2(a) of criteria elements, the values as applicable within coastal and territorial waters;
(b) for contaminants selected under point 2(b) of criteria elements, the concentrations for a specified matrix (water, sediment or biota) which may give rise to pollution effects. Member States shall establish these concentrations through regional or subregional cooperation.
[2] Diclofenac is tested in both seawater and biota but neither matrix has been designated as the primary testing matrix. For the SOM analysis, seawater is used due to the better geographic coverage of samples.

The GES threshold for PFOS is set at 9.1 µg/kg fish muscle wet weight in the primary matrix fish muscle. In the latest HOLAS assessment period (2011-2016), PFOS was in a not good status at the Baltic Sea scale. However, this was the result of not good status in a single coastal area and consistent not good status in the secondary matrix of seawater, where sampled (HELCOM 2018a, 2018c). The SOM analysis only considers the status as assessed in the primary matrix and therefore may be in conflict with the latest status assessment. For TBT, the primary matrix is sediment with a test threshold of 1.6 µg/kg dw sediment (5% TOC). In the latest HOLAS assessment period (2011-2016), TBT fails to achieve good status at both the Baltic scale and in all surveyed sub-areas (HELCOM 2018a, 2018d). Several secondary matrices exist, which also indicate widespread not good status. The test threshold for diclofenac is 0.005 µg/l sea water. Diclofenac is also tested in biota and neither matrix has been designated as the primary testing matrix. For the SOM analysis, seawater is used due to the better geographic coverage of samples. In the latest HOLAS assessment period (2011-2016), diclofenac was not quantitatively assessed (HELCOM 2018a). However, concentrations failing to reach good state are generally restricted to within 10 km of wastewater treatment plant outfalls. An optimal monitoring program is still under development (HELCOM 2018e).

Hazardous substances are also considered in the SOM analysis as more general pressures on the Baltic Sea environment. Six general pollution pressures related to hazardous substances could be selected when identifying the most significant pressures linked to any of the state components included in the SOM analysis: *Heavy metal pollution, Organohalogen pollution, Organotin pollution, Pharmaceutical pollution, Hydrocarbon pollution, Radioactive pollution* (Figure 1). These pressures are more broadly defined than the HELCOM indicators but do correspond with the MSFD criteria D8C1[1]. These pressures are included in the SOM analysis to capture the overall effects of hazardous substances to the environment, not limit the analysis to the four substances and to accommodate the varying knowledge of hazardous substances of experts in other fields, e.g. marine mammals, waterbirds, etc. Management and monitoring of hazardous substances require the substance specific approach taken for mercury, PFOS, TBT and diclofenac, but accounting for the overall impact of hazardous substances in the SOM model does not.

Supplementary activities

As supplemental information to the update of the BSAP and SOM analysis, background documents were prepared on each of the target substances and their sources and pathways as part of the SOM work. Background documents were additionally prepared for PBDEs, dioxins and PCBs, as these are substance groups also currently deteriorating the status. These substance groups were not considered in the SOM model due to limited resources. Links to these all the background reports are available below.

- Diclofenac
- PFOS and other PFASs
- Dioxins and PCBs
- PBDEs
- Mercury (link to be added later)
- TBT (link to be added later)

Methods and data

The section below includes an overview of any topic-specific methodologies. A full description of the general approach, methods and data collection for the SOM analysis is available in this document. Note that the detailed results are presented for the most likely development of human activities and using the expert data on effectiveness of measures.

Activity-pressure input contributions

The contributions of activities to the input of each four substances were determined using surveys that were distributed to national topic experts via the HELCOM Expert Network on Hazardous Substances (EN HZ). Responses from individual experts were accepted, but because national responses were preferred, all responses were weighted nationally to standardize the data set. The background documents on the target substances were provided as supplementary material to the experts to support their survey response. Respondents were asked to assess the maximum, minimum, and most likely contribution of any activity contributing more than 5% to the total input of each substance. Responses to activities contributing below that threshold were invited but not required. Respondents were also asked to assess the extent to which data informed their answer using a five-point scale (1 being very low and 5 very high).

Effectiveness of measures and pressure-state linkages

Measure types (Annex 3) and structural relationships between the measure types and activities and pressure inputs (Annex 7) were designed by the HELCOM Workshop on the analysis of sufficiency of measures for hazardous substances (HELCOM SOM-HZ WS 1-2019) in collaboration with the Hazardous Substances Topic Team and HELCOM ACTION WP6. The measure types were informed by the existing measures list (Annex 4), but were also designed to acknowledge the full breadth of potential measures.

For hazardous substances, the effectiveness of measures survey structure comprised 37 unique measure types covering 11 activities. The same measure type may be listed under multiple activities and pressure inputs. Altogether this resulted in 45 assessments of measure type effectiveness across the four pressure inputs, *Input of mercury, Input of PFOS, Input of TBT,* and *Input of diclofenac*. The exact list of measure types, and their grouping by activities and pressure inputs is shown in Annex 7. The effectiveness of measures survey itself is included as Annex 8.

Effectiveness of the measure types and links between the pressures and state components were determined using online expert surveys implemented in December 2019 – February 2020 with follow-up surveys conducted in the spring 2020. The expert pool consisted of the HELCOM Expert Network on Hazardous Substances, the HELCOM Correspondent Group on Pharmaceuticals, participants of the HELCOM Workshop on the analysis of sufficiency of measures for hazardous substances and nationally nominated experts. Additionally, the project received survey responses from experts not on the original invitation list; these responses were also included in the analysis. The full description of the methodology and data collection is available as part of the SOM methodology report.

Topic specific model structure, assumptions and challenges

The SOM analysis for hazardous substances evaluates the sufficiency of measures in achieving GES, considering the effects of existing measures and future development of human activities. In the SOM model structure, the pressure inputs *Input of mercury, Input of PFOS, Input of TBT, Input of diclofenac* are assumed to be equivalent to the relevant SOM pressures *Heavy metal pollution, Organohalogen pollution, Organotin pollution, Pharmaceutical pollution* for the analysis of specific substances. In practice this means that e.g. *Heavy metal pollution* is considered to represent only mercury pollution for the assessment of the *Concentration of mercury*. The assumption relies on absent or limited interactions within a class of hazardous substances e.g. the input of other heavy metals (Pb, Cd, etc.) does not lead to an increase in mercury, the substance of interest. This assumption is not made when linking the hazardous substance pressures to any other state variable.

Existing concentrations of the target substances are not directly considered in the analysis. When assessing the required pressure reduction to reach good status, experts were provided links to assessments of current status and expert responses regarding time lags likely considered current status as well. This approach could have difficulties with topics with long time lags, such as exists for several hazardous substances. The SOM model is static and thus is not ideal for capturing the complex temporal dynamics present in e.g. burial and resuspension. Use of a dynamic model would improve coverage of these aspects but would be too complex to currently implement.

This challenge also affects the ability for topic experts to provide an accurate required pressure reduction to reach GES. For a topic with significant time lags to recovery, concentrations would be expected to continue to fall for years or decades even if no additional steps were taken. When compiling the list of existing measures, countries were asked to include measures implemented prior to the base year (2016) that had time lags between measure implementation and pressure reduction that would indicate additional impact occurring after the start of the base year (2016 forward). This mechanism allows measure-pressure input time lags to be considered in the analysis. However, no such mechanism exists for the time lags between pressure and state (time lags to recovery). The difficulty of correctly estimating the required pressure reduction to reach GES increases with substance time lags to recovery and in these cases more interpretive weight must be given to the estimated time lags.

Overview of data

Table 1 shows the origin and spatial resolution for the data components in the SOM analysis for hazardous substances. Activity-pressure input contributions, pressure-state links and time lags are based on expert data. Information on existing measures comes from literature reviews and Contracting Parties, and development of human activities is based on existing literature, data and projections.

Estimates of the effectiveness of measures were collected both via expert surveys and a literature review for all topics included in the SOM analysis. The aim of the literature review was to compile information from scientific articles and reports providing estimates on the effects of measures in reducing pressure inputs that could be used in the SOM analysis, either by including the estimates in the SOM model or by providing comparison points. The literature review was conducted by topic, with the information collected into structured excel files (see the methodology document, Annex 5 and Annex 6 for more information). For hazardous substances, 134 effectiveness estimates from 44 studies were compiled. Out of these, 72 estimates from 23 studies could be included in the model. Detailed results are presented using only the expert data, and the implications of using the literature data for the effectiveness of measures are reviewed in the discussion section.

The spatial resolution (level of detail) differs across the data components of the SOM analysis. All assessment areas are based on the 17 HELCOM scale 2 sub-basins and the assessment area ranges from the single Baltic Sea to individual sub-basins. However, for hazardous substances, all expert data are reported at the Baltic Sea scale, while data on existing measures and their implementation status exists at the sub-basin level. When the topic of hazardous substances interacts with other topics, e.g. birds, mammals and benthic habitats, smaller spatial scales may be used to reflect spatial variation in those topics. Table 1 shows the origin and spatial resolution for the data components in the SOM analysis for hazardous substances.

Table 1. Data for hazardous substances (more information on data collection is available in the methodology document).

Data component	Origin of data	Spatial resolution
Activity-pressure input contributions	Expert evaluation	Whole Baltic Sea
Existing measures	Literature review, Contracting Parties	17 sub-basins
Effectiveness of measures	Expert evaluation, literature review	Whole Baltic Sea
Development of human activities	Literature review, existing data and projections	Whole Baltic Sea
Pressure-state links	Expert evaluation	Whole Baltic Sea
Time lags	Expert evaluation	Whole Baltic Sea

Development of human activities

In addition to existing measures, changes in the extent of human activities may affect pressure inputs over time. Four scenarios for future changes in human activities were developed: 1) no change, 2) low change, 3) moderate (most likely) change, and 4) high change. These alternative scenarios aim to capture uncertainties and variation in the future development of human activities. The results of the SOM analysis were estimated for each of the four scenarios to assess how the alternative assumptions on the development of human activities affects the findings. Detailed results are presented for the most likely development scenario, and implications of using the other scenarios on the results are reviewed in the discussion section.

The scenarios specify a percent change in each activity in 2016–2030 based on existing information and projections from the Baltic Sea region. Change scenarios were made only for predominant activities in the Baltic Sea region, including agriculture, forestry, waste waters, (commercial) fish and shellfish harvesting, aquaculture, renewable energy production, tourism and leisure activities, transport shipping and transport infrastructure. Other activities are assumed to stay unchanged. This means that only 9 of the 31 standard SOM activities have change scenarios in the SOM analysis. This results in varying influence of these scenarios on the results across topics, pressures and state components, depending on the significance of the activities to the pressure inputs relevant to the topic.

For hazardous substances, coverage of activities that contribute to pressure inputs in the change scenarios is very low for mercury (5%), low for PFOS (22%) and TBT (26%), and high for diclofenac (56%). Important activities contributing to the input of these substances that are lacking development scenarios include non-renewable energy generation (mercury), air transport (PFOS), industrial uses (mercury), solid waste (PFOS; diclofenac), restructuring of seabed morphology (TBT), and activities and sources outside the Baltic Sea region (mercury).

The current situation with COVID-19 and its possible implications to the development of human activities is not reflected in the scenarios, as there is no information on the long-term effects it may have on the economy or activities. The current situation poses a challenge for choosing the most likely scenarios for the development of human activities, which has been done based on currently available information.

Results and interpretation

Background

The SOM results are presented in the format of percent shares or probabilities. The main finding of the analysis is the probability to achieve GES or specific state improvements/pressure input reductions, taking into consideration the effects of existing measures and changes in the activities on pressure inputs. The contribution of activities to pressure inputs, the effect of measures on pressure inputs, and the significance of pressures to state components are presented as percent values (e.g. how many percent would the measure reduce the pressure input). Results are presented mainly in tables, which show the the most likely (expected) values and standard deviations. Standard deviation is a way of showing the variation in the values. When it is high, values are spread over a wider range, and when it is low, values are closer to the most likely value. Figures and graphs presenting distributions are included in the annexes. They show the same results as the tables but allow either more detailed information or alternative visualisation of the results.

For the data that are based on expert surveys, the confidence rating gives the most common answer to experts' assessment of the confidence in their own survey responses on a low-moderate-high scale. More detailed information on how each result has been calculated is presented in a separate document.

This document presents the detailed results based on the expert-based data (survey responses). Literature data on the effectiveness of measures has been collected and included in an alternative model estimation. The impacts of using the literature data are evaluated in the discussion section. In the detailed results, the projected development of human activities is based on the most likely future development until 2030 (for details, see the methodology document), and the impacts of alternative scenarios on human activities are examined in the discussion section.

Format of presentation

The format the results are reported in (not presented, qualitative/semi-quantitative, quantitative) depends on the type of result and the number of participating experts. Further, for all results utilizing other SOM results as input data, reporting is done at the most conservative standard used in the input data. In practice this means that if one input data point is reported as ‘insufficient data’, all results using that data point will also be reported as ‘insufficient data’; similarly for qualitative/semi-quantitative data points. However, note that this standard is only applied in the case of data points actively used to calculate another result. For example, many measure types are hypothetical or otherwise not implemented in the Baltic Sea and therefore do not factor into results on projected pressure input reductions from existing measures. Insufficient data for such measure types does not affect reporting other results that rely on data for effectiveness of measure types. Results that do not meet the data standards described here and in greater detail below are marked with ‘insufficient data’ in the report. All the data components for hazardous substances meet the thresholds for fully quantitative presentation.

For results concerning required pressure reductions and significance of pressures to state components, results with 2 or fewer respondents are not reported; results with 3 to 4 respondents will be either not reported, or qualitatively/semi-quantitatively reported based on feedback from the SOM topic teams or other HELCOM expert body; results with 5 or more respondents are reported quantitatively. This standard allows flexibility for reporting on assessments that are of spatially limited areas and therefore have fewer experts available to survey, while also being somewhat conservative in reporting fully quantitative results.

For expert-based effectiveness of measures results, measure types with 5 or more respondents are reported quantitatively and those with 4 or fewer respondents are listed as having insufficient data.

For expert-based activity-pressure input results, expert responses where primarily sought through the HELCOM expert networks in the form of national responses. Individual expert responses were accepted but were consolidated into average responses by country to conform to the format of other responses. Thus, the maximum number of responses is 9. This maximum is rarely reached due to responses typically only applying to areas adjacent to the specific country. Acknowledging this, activity-pressure input relationships are reported if there are expert responses from 3 or more countries or if the number of countries providing expert responses is greater than 1/2 the number of countries bordering any given sub-area (see Table 2 below; responses from experts based in any HELCOM country will be counted toward the reporting threshold, i.e. the reporting assessment is not limited to responses from bordering countries).

Table 2. Required number of countries providing expert responses to the activity-pressure input survey to meet the minimum data threshold for reporting.

Bordering countries	Required number of countries providing expert responses to meet minimum data threshold	Example areas
1	1	Western Gotland Basin
2	2	Bothnian Sea, Gulf of Riga
3	2	Gulf of Finland
4+	3	Eastern Gotland Basin, Baltic Sea

Coverage of pressures in the SOM analysis

The SOM analysis has only been able to account for a portion of all pressures that affect the state components, and the effect of several significant pressures have not been included due to not being able to quantify the link between the pressure inputs, pressures and state components in the analysis. This means that the effect of reductions in these excluded pressures on the state components is not included in the total pressure reductions, and the projected total pressure reductions and probability to achieve GES are underestimated. The share of pressures covered in the analysis has been calculated based on the significance of pressures to the state component in question. The share varies across topics and state components from low (around 20%) to high (more than 80%).

Are existing measures sufficient for achieving good status?

Concentrations of mercury and PFOS have established HELCOM GES thresholds and concentrations of TBT and diclofenac are currently evaluated by applying preliminary GES thresholds. Thus, it is possible to assess whether existing measures are sufficient in achieving GES.

Overall, the results of the analysis indicate that existing measures are not sufficient in achieving GES for the concentrations of mercury, TBT, PFOS and diclofenac. Reductions in concentrations are around 10-30% for mercury, around 5-30% for PFOS and close to zero for TBT and diclofenac (Table 3). Required pressure reductions to achieve GES are in the range of 20-85% (Table 4), thus indicating considerable uncertainty about how much pressures need to be reduced to achieve GES. The probability to achieve GES with expected total pressure reductions is very low for mercury, TBT and diclofenac, and low for PFOS.

In the latest HOLAS assessment period (2011-2016), mercury was in a not good status at the Baltic Sea scale; with good status achieved only in a few smaller scale areas in the southwestern Baltic (HELCOM 2018a-b). PFOS was in a not good status at the Baltic Sea scale, however, this was the result of not good status in a single coastal area and consistent not good status in the secondary matrix of seawater, where sampled (HELCOM 2018a, 2018c). TBT failed to achieve good status at both the Baltic scale and in all surveyed sub-areas (HELCOM 2018a, 2018d). Diclofenac was not quantitatively assessed (HELCOM 2018a).

The near zero projections of the probability of achieving GES for TBT and diclofenac are due to few or no existing measures to control their inputs (or release from historic deposits), and aspects such as natural processes (e.g. sediment burial and retention) are not explicitly included in this evaluation. The result from PFOS is unexpected due to widespread good status of PFOS through the Baltic Sea in 2016 (37/38 areas in good state; HELCOM 2018c) and strong control of the substance under the Stockholm Convention on Persistent Organic Pollutants. This could potentially be a result of experts evaluating the state of PFOS for all potential testing matrices rather than for the primary matrix of fish muscle. The secondary matrix seawater fails to achieve good status in all tested areas (0/8 areas in good status; HELCOM 2018c). Alternatively, this may reflect concerns about future input from legacy sources or broader concern about PFAS generally.

In the case of mercury, TBT, PFOS and diclofenac, the SOM analysis has been able to account for 61-85% of the pressures linked to the substances (pressures highlighted in white in Table 5). This percent reflects the share of pressures that 1) have a quantifiable link to the concentrations of the four substances and 2) have measures types that affect them in the SOM analysis. It has been calculated based on the significance of pressures to the concentrations of the four substances. It is the maximum pressure reduction that could be achieved if the pressures linked to the concentration of the four substances in the SOM analysis were eliminated. The effects of several significant pressures are not included in this total, such as eutrophication and change in hydrologic conditions (pressures highlighted in grey in Table 5). Although these pressures are expected to decrease based on the results of the SOM analysis, the analysis is not able to estimate how this would affect the concentrations of hazardous substances. Thus, the total pressure reductions and probability to achieve GES are underestimated.

The assessment compares the state improvement from existing measures to the state improvement required to achieve GES. The calculation of sufficiency of measures takes into account all the components of the SOM analysis: the activity-pressure input contributions, effectiveness of measure types in reducing pressure inputs, links between existing measures and measure types, projected pressure input reductions from existing measures, development of human activities, significance of pressures to state components and pressure reductions required to achieve GES. The analysis assumes that all existing measures are fully implemented and that there are no time lags between the input of hazardous substances and their concentrations.

The sufficiency of measures result is presented as the probability of achieving GES with the projected total reduction in pressures by substance at the scale of the entire Baltic Sea. Table 3 shows the expected total pressure reductions from existing measures, the probability to achieve GES with such a pressure reduction, and the maximum pressure reduction that could be achieved with the fully quantified pressures in the SOM analysis. Total pressure reductions are calculated based on the reduction in the input of mercury, TBT, PFOS and diclofenac (Table 7), significance of different pressures to the concentrations of these substances (Table 6), and spatial weighting to account for the target area of existing measures.

Table 4 shows the average of the mostly likely total pressure reduction required to reach GES for each substance, based on the expert responses. There is considerable uncertainty among experts about the required pressure reductions to achieve GES, as the standard deviations are high compared to the most likely value. The required pressure reductions are 60-75% for mercury, 35-75% for TBT, 35-65% for PFOS and 30-60% for diclofenac, based on the 90% confidence intervals. This indicates that substantial pressure reductions would be required to achieve GES, but their magnitude is very uncertain. The experts' confidence in their own estimates is moderate.

Distributions of expert responses on the required pressure reductions to achieve GES are included in Annex 10. Experts have differing opinions and there is substantial uncertainty about the pressure reductions required to achieve GES. For example, in the case of mercury, experts seem to agree that pressure reductions that are lower than 40% would not be sufficient to achieve GES, but have varying views on the size of the pressure reductions needed. Thus, there is considerable uncertainty about the link between pressure reductions and achieving GES for these four hazardous substances.

[Further discussion and interpretation of the main results, input from topic team]

Table 3. Sufficiency of measures in achieving GES for concentrations of mercury, TBT, PFOS and diclofenac in the Baltic Sea. The table presents the expected values and the 10-90 percentile in brackets, which shows the range in which 80% of the observations fall in.

State	**Assessment area**	**Total pressure reduction (%)** [10 percentile – 90 percentile]	**Probability to achieve GES (%) with expected pressure reduction** [10 percentile – 90 percentile]	**Maximum possible pressure reduction due to model coverage (%)**
Mercury concentration	Baltic Sea	20 [10 - 29]	0 [0-0]	61
TBT concentration	Baltic Sea	-1 [-12 - 7]	0 [0-0]	75
PFOS concentration	Baltic Sea	16 [7 - 27]	13 [1 - 13]	68
Diclofenac concentration	Baltic Sea	0 [-3 - 2]	0 [0 - 1]	85

Data used: activity-pressure input contributions, effectiveness of measure types, information on existing measures, significance of pressures to state components, required pressure reductions to achieve GES, development of human activities

Table 4. Total pressure reduction required to reach GES. Standard deviation is given in parentheses. Values are calculated directly from expert survey data. Confidence depicts the most common rating of expert's confidence in their own responses to the question on total pressure reduction required to reach GES.

State	Mercury concentration	TBT concentration	PFOS concentration	Diclofenac concentration
Most likely pressure reduction required (%)	69 (15) ●●●	55 (27) ○●●	51 (21) ○●●	46 (24) ○●●
Confidence	Moderate	Moderate	Moderate	Moderate
Number of experts	11	8	8	11

Colour scale for the percent reduction in pressures required to reach GES in percent (based on the expected value): 0-10%, 10-20%, 20-40%, 40-60%, 60-100%

Categories for the certainty of the reduction required estimate (based on the relative size of the standard deviation to the expected value): low: ○○●, moderate: ○●●, high: ●●●

Data used: required pressure reductions to achieve GES

What are the time lags between pressure and state?

Information on time lags between reducing the pressure inputs and concentrations was collected from experts, who evaluated how long it would take to achieve GES assuming sufficient measures were implemented. Table 5 shows the distribution and average of the answers for the four substances.

The likely time lag for mercury is considered to be the longest out of the substances, with an average of 38 years. Achieving GES for TBT and PFOS with sufficient measures is evaluated to take on average 14 and 17 years, respectively, while time lag for diclofenac is considerably smaller, on average 5 years. These expert evaluations indicate that even with sufficient measures, it takes time to achieve GES concentrations in the Baltic Sea for the selected (and likely other) hazardous substances.

The main reported factors contributing to the time lag for mercury were high retention in sediments leading to long-term release into the Baltic Sea, slow burial rates, and that it is non-degradable. The primary reported factors for TBT were its long degradation time and local sedimentation rates. Degradation time and sedimentation rates were also reported for PFOS, in addition to the long residence time of water in the Baltic Sea. The primary reported factor influencing diclofenac was its degradation time.

Additional information on time lags related to hazardous substances will be produced in HELCOM ACTION project WP5.

[Further discussion and interpretation, input from topic team]

Table 5. Time lags in achieving GES with sufficient measures. Responses with clear reference to time lags due to lags in the implementation of measures have been excluded.

Time lag	Mercury	TBT	PFOS	Diclofenac
0 years (no time lag)	0	0	0	0
0-5 years	0	0	1	6
6-10 years	1	4	1	1
11-25 years	4	1	3	1
26-50 years	3	1	1	0
51-100 years	3	0	0	0
More than 100 years	0	0	0	0
Excluded	0	2	1	3
Average	37.7	14.2	16.7	5.0
SD	24.7	11.1	11.0	5.0
Confidence	1.9	2.0	2.2	2.1
Number of experts	11	6	6	8

Data used: expert responses on time lags

What are the pressures contributing to the state components?

This section presents the significance of different pressures to the concentrations of mercury, TBT, PFOS and diclofenac based on the responses to the expert surveys. They are all assessed at the Baltic Sea scale and enable comparisons across substances. Based on the responses, the number of pressures affecting the concentration varies from three (diclofenac) to nine (mercury) (Table 6). Confidence in Table 6 depicts the most common rating of expert's confidence in their own responses to the significance of pressures question, and it is high for most substances.

As expected, the most significant pressure is in all cases the pollution associated with the specific hazardous substance in question, with significance of 45-80%, but also several other pressures are considered to affect the concentrations. Change in hydrologic conditions is evaluated to be significant for the concentrations of all substances, and physical disturbance of marine habitats is particularly significant for TBT concentration.

[Further discussion and interpretation, input from topic team]

Table 6. Significance of pressures (%) affecting the concentration of mercury, TBT, PFOS and diclofenac in the Baltic Sea.

State component **Pressure**	**Mercury concentration**	**TBT concentration**	**PFOS concentration**	**Diclofenac concentration**
Effects of non-indigenous species		7		
Physical disturbance of marine habitats		22	2	4
Physical loss of marine habitats	8			
Effects of marine litter (excluding bycatch in ghost nets)			9	
Effects of eutrophication	3	2		
River, lake, or land habitat loss/degradation	6	2	13	
Hydrocarbon pollution	2			
Radioactive pollution	4			
Organohalogen pollution (e.g. PFOS, PCBs, PBDEs, dioxins)			66	
Organotin pollution (e.g. TBT)		47		
Heavy metal pollution	48	5		
Pharmaceutical pollution	6			81
Change in hydrologic conditions	19	15	11	15
Human-induced food web imbalance	4			
Confidence	High	High	High/ moderate	High
Number of experts	11	7	7	9

Colour scale for the significance of the pressure to the state variable (based on the expected value):
0-10%, 10-20%, 20-40%, 40-60%, 60-100%
Pressures for which we cannot quantify the link between the pressure input, pressure and state in the SOM analysis are highlighted in grey, e.g. we cannot link reductions in nutrient inputs to reductions in the effects of eutrophication and further to concentrations of hazardous substances.
Data used: expert responses on significance of pressures to state components

What are the state components most affected by hazardous substances?

The expert data allow for listing the state components most affected by the pressure in question. This is based on data from all five pressure-state expert surveys, where significance of pressures to state components was assessed.

Table 7 shows the state components most affected by the pollution pressures related to hazardous substances. Each of the pressures related to the four substances (heavy metal pollution, organotin pollution, organohalogen pollution, pharmaceutical pollution) affects the most the concentration of the relevant targeted substance (mercury, TBT, PFOS, diclofenac). Other state components affected are concentrations of other hazardous substances, bird species, fish species, grey seal, and benthic habitats.

[Further discussion and interpretation, input from topic team]

Table 7. Top five state components most affected by hazardous substance pollution. Listing is based on Baltic-wide averages of the significance of pressures to state components presented in each respective topic report. Average number of expert responses for the state component is given in parenthesis (total response count for the state component divided by the number of geographic areas for the state component).

Pressure	1st most affected state component	2nd most affected state component	3rd most affected state component	4th most affected state component	5th most affected state component
Hydrocarbon pollution	Long-tailed duck (7)	Red-throated diver (6)	Great black-backed gull (5)	Grey seal (5)	Mercury concentration (11)
Radioactive pollution	Mercury concentration (11)	Cod (15)	Hard substrate vegetation dominated community (5.8)	Soft substrate vegetation dominated community (3.8)	Hard substrate epifauna dominated community (5.3)
Organohalogen pollution (e.g. PFOS, PCBs, PBDEs, dioxins)	PFOS concentration (7)	Grey seal (5)	Insufficient data (less than 3)	Insufficient data (less than 5)	Eel (11)
Organotin pollution (e.g. TBT)	TBT concentration (7)	Hard substrate vegetation dominated community (5.8)	Soft substrate vegetation dominated community (3.8)	Hard substrate epifauna dominated community (5.3)	Soft substrate infauna dominated community (5)
Heavy metal pollution	Mercury concentration (11)	Insufficient data (less than 3)	Soft substrate infauna dominated community (5)	Eel (11)	TBT concentration (7)
Pharmaceutical pollution	Diclofenac concentration (9)	Mercury concentration (11)	Soft substrate infauna dominated community (5)	Hard substrate vegetation dominated community (5.8)	Soft substrate vegetation dominated community (3.8)

Data used: expert responses on significance of pressures to state components for all topics
'Insufficient data' indicates cases where there are not enough responses to the significance of pressures to the state component in the expert survey (e.g. some mammals species), corresponding to the criteria for the format of presentation.

What are the reductions in pressure inputs from existing measures?

Table 8 shows the effects of existing measures in reducing the input of hazardous substances to the Baltic Sea in 2016-2030, considering the changes in the extent of human activities. They are calculated using the data on activity-pressure input contributions, effectiveness of measure types, links between existing measures and measure types, and projected development of human activities.

The activity-pressure input and the effectiveness of measures data are at the Baltic Sea level, and thus the total pressure reductions are presented as an average for the entire Baltic Sea. The projected reductions in pressure inputs account for the joint impacts across the measure types, as well as the spatial area where the pressure inputs can be reduced to avoid overestimating the pressure input reductions. Pressure input reductions can be positive, negative or zero, depending on the combined effect of existing measures and changes in the extent of human activities. When the reduction in pressure inputs from existing measures is larger than the increase from changes in human activities, pressure inputs are reduced.

Of the four hazardous substances, the *input of mercury* is expected to be reduced the most, from a moderate to high extent. The main activities contributing to the input of mercury are expected to remain constant until 2030. *Input of PFOS* is expected to decrease moderately. Out of the main activities, changes are expected only in waste waters, where minor increases are expected to urban sewage systems.

Negative projected reduction for the *input of TBT* implies that the input is projected to increase. This happens when the pressure input reductions from existing measures are unable to compensate for the pressure input increases caused by the projected future development of activities. For TBT, measures to control inputs (other than a ban on use) are not applied in the Baltic Sea, and therefore this result is influenced only by the projected development of activities. Increases in the extent of shipping and transport infrastructure and tourism and leisure activities lead to increases in the input of TBT. These may increase e.g. release from existing sources or deposits.

The *input of diclofenac* is expected to stay the same. Existing measures are implemented in the Bay of Mecklenburg, Arkona Basin, and Bornholm Basin. In the other sub-basins, projected future changes in the activities lead to increased inputs of diclofenac, e.g. from based on increases in urban waste waters, which is a significant contributor to the input of diclofenac.

Overall, there is rather high uncertainty about the projected reductions in the input of hazardous substances, as shown by the large standard deviations. This stems from the uncertainty on both the effectiveness of measure types and activity-pressure input contributions. The certainty of the estimates is evaluated as low for TBT and diclofenac, and moderate for mercury and PFOS.

The impact of future development in the extent of human activities to the input of the substances is limited, as the main activities contributing to the input of the substances are generally assumed to remain constant (e.g. non-renewable energy generation, solid waste, restructuring of seabed morphology, activities and sources outside the Baltic Sea region). Further details on the effectiveness of different measure types and activity-pressure input contributions can be found in Tables 9 and 10.

[Further discussion and interpretation of the results, input from topic team]

Table 8. Projected reductions (%) in the input of hazardous substances from existing measures in the Baltic Sea. The table depicts the most likely/expected values of reductions in pressure inputs and gives standard deviations in parenthesis.

Pressure input Area	Input of mercury	Input of TBT	Input of PFOS	Input of diclofenac
Baltic Sea	38 (15) ○●●	-13 (15) ○○●	24 (11) ○●●	-2 (2) ○○●

Colour scale for the pressure input reductions in percent (based on the expected value):
<0%, 0-10%, 10-20%, 20-40%, 40-60%, 60-100%
Categories for the certainty of the pressure input reductions (based on the relative size of the standard deviation to the expected value): low: ○○●, moderate: ○●●, high: ●●●
Data used: activity-pressure input contributions, effectiveness of measure types, information on existing measures

How effective are measure types in reducing pressure inputs?

This section presents the percent effectiveness of measure types in reducing the input of mercury, TBT, PFOS and diclofenac from a specific activity. The estimates are presented per activity, i.e. they portray the percent reduction in the pressure input from the activity in question, and not in the total input across all activities. Information on the reductions over all activities contributing to the pressure input is given in the section on the impacts of measure types. Data on the effectiveness of measure types originate from expert surveys and are at the Baltic Sea scale.

In the following, percent effectiveness is presented per pressure input, activity and measure type, and pooled over experts. The effectiveness estimates can be compared across measure types to assess, on average, how effective they are in relation to each other in reducing the pressure input from the specific activities, or across activities to assess which measure type could be the most effective for each activity.

Tables 9.1 – 9.4 present the expected percent effectiveness of the measures type and its standard deviation. Confidence in Tables 9.1 – 9.4 depicts the most common rating of expert's confidence in their own responses to the effectiveness of measure types question. Annex 11 presents the distributions of the effectiveness of measure types in controlling the input of the four substances for additional information.

For the input of mercury, each measure type can only reduce the pressure input from a single activity (Table 9.1). The measure types of *local/state/national targets to eliminate coal fired energy production* for non-renewable energy generation and *treatment of scrubbing water from ships before disposal to reduce mercury* for shipping have the highest effectiveness for that activity. The effectiveness of the rest of the measure types ranges between 20-60% in reducing the input from the specific activity.

Table 9.2 shows the effectiveness of measure types to in reducing the input of TBT from three activities. Each measure type can only reduce the pressure input from one activity. The most effective measures types are *reducing re-suspension from sediments, by limiting restructuring of seabed to areas with low concentrations* for activity restructuring of seabed morphology, *boat washing restrictions* for tourism and leisure activities and *in water hull cleaning regulation* for shipping.

Table 9.3 shows the effectiveness of measure types in reducing the input of PFOS from five activities. The input from activities *transport – air* and *urban uses (land use)* can only be reduced by a single measure type. *Clean-up of contaminated sites* is the most effective measure type in reducing the pressure input from industrial uses and solid waste. *Implementing technologies to remove PFOS from wastewater* is the most effective measure type for reducing the input from waste waters.

Table 9.4 shows that the input of diclofenac is mainly influenced by the activity *waste waters*. The measure types *technical upgrade of wastewater treatment plants* and *alter prescription practices to lower consumption* are the most effective ones.

Overall, there is considerable uncertainty about the effectiveness of certain measures types base on the standard deviations. The certainty of the estimates varies from low to moderate. Confidence of the estimates is high for measure types affecting diclofenac, and most often moderate for mercury, TBT and PFOS.

Estimates of the effectiveness of measure types are used to assess the effects of existing measures in reducing the input of hazardous substances to the Baltic Sea and to calculate pressure input reductions from existing measures by 2030.

[Further discussion and interpretation of results, input from topic team]

Table 9.1 Effectiveness of measure types (%) in reducing the potential *input of mercury.* The effectiveness of a measure type is the percent reduction in the pressure input resulting from a specific activity. The table depicts the expected effectiveness, and standard deviation is given in parenthesis.

Measure type ID	Activity Measure type	Non-renewable energy generation	Transport – land	Industrial uses	Waste waters	Solid waste	Restructuring of seabed morphology	Transport – shipping	Activities and sources outside the Baltic Sea Region	Has corresponding existing measures in the SOM analysis (Yes/No)
1	Local/state/national targets to eliminate coal fired energy production	70 (29) ○●●	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Yes
2	Continued application of the EU Emissions Trading System	34 (20) ○●●	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Yes
4	Sand capping or removal of existing cellulose deposits	Not assessed	Not assessed	30 (23) ○○●	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	No
5	Ban on manufacturing processes where mercury or mercury compounds are used as a catalyst (e.g. In vinyl chloride monomer production	Not assessed	Not assessed	22 (23) ○○●	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Yes
6	Ban on manufacturing processes where mercury is used as an electrode (e.g. in chlor-alkali production; Na or K methylate/ethylate production	Not assessed	Not assessed	28 (22) ○○●	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Yes
7	Ban on export of Mercury (II) sulphate (HgSO4, CAS RN 7783-35-9) and Mercury (II) nitrate (Hg(NO3)2, CAS RN 10045-94-0)	Not assessed	Not assessed	22 (23) ○○●	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Yes
8	Treatment of scrubbing water from ships before disposal to reduce mercury	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	63 (32) ○●●	Not assessed	Yes
9	Reduce re-suspension from sediments, by limiting restructuring of seabed to areas with low concentrations	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	51 (26) ○●●	Not assessed	Not assessed	No

Measure type ID	Activity Measure type	Non-renewable energy generation	Transport – land	Industrial uses	Waste waters	Solid waste	Restructuring of seabed morphology	Transport – shipping	Activities and sources outside the Baltic Sea Region	Has corresponding existing measures in the SOM analysis (Yes/No)
10	Tighter allowed contamination levels during dredging activities	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	43 (24) ○●●	Not assessed	Not assessed	No
11	Perform dredging under conditions (low water pH and temperature) that lower desorption from sediments, i.e. Preferable during winter	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	27 (22) ○○●	Not assessed	Not assessed	No
12	Increased fuel efficiency standards	Not assessed	35 (24) ○○●	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	No
13	Increased electrification of transportation fleets	Not assessed	41 (25) ○●●	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	No
14	Minamata convention	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	47 (18) ○●●	Yes
15	Paris agreement	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed	44 (20) ○●●	Yes
16	EU mandatory use of dental amalgam separators retaining at least 95% of amalgam particles	Not assessed	Not assessed	Not assessed	40 (24) ○○●	Not assessed	Not assessed	Not assessed	Not assessed	Yes
17	5% increase in EU mandatory WEEE recycling levels	Not assessed	Not assessed	Not assessed	Not assessed	36 (32) ○○●	Not assessed	Not assessed	Not assessed	Yes
18	Further restrictions on storage and disposal of waste/dredged material containing mercury	Not assessed	Not assessed	Not assessed	Not assessed	44 (30) ○○●	Not assessed	Not assessed	Not assessed	Yes
140	Eurasian Economic Union restrictions on hazardous substances in electrical products	Not assessed	Not assessed	Not assessed	Not assessed	31 (21) ○○●	Not assessed	Not assessed	Not assessed	Yes
	Confidence	Moderate	Low-High	Low	Moderate-High	Moderate	Moderate	Moderate-High	Moderate	
	Number of experts	7	7	6-7	7	5-6	7	7	7	

Colour scale for the effectiveness of a measure type in percent (based on the expected value): 0-10%, 10-20%, 20-40%, 40-60%, 60-100%

Categories for the certainty of the effectiveness estimate (based on the relative size of the standard deviation to the expected value): low: ○○●, moderate: ○●●, high: ●●●
Data used: expert responses on effectiveness of measure types
Full activity names:

- Non-renewable energy generation (fossil fuel and nuclear powerplants)
- Transport – land (cars and trucks, trains), including infrastructure
- Industrial uses (oil, gas, industrial plants)
- Waste waters (urban, industrial, and industrial animal farms; includes all waste streams entering wastewater systems e.g. microplastics, pharmaceuticals, etc.)
- Solid waste (e.g. land-based disposal of dredged material, landfill, solid waste streams)
- Restructuring of seabed morphology (dredging, beach replenishment, sea-based deposit of dredged material)
- Transport – shipping (incl. anchoring, mooring)
- Activities and sources outside the Baltic Sea Region

Table 9.2 Effectiveness of measure types (%) in reducing the potential *input of TBT.* The effectiveness of a measure type is the percent reduction in the pressure input resulting from a specific activity. The table depicts the expected effectiveness, and standard deviation is given in parenthesis.

Measure type ID	Activity Measure type	Restructuring of seabed morphology	Tourism and leisure activities	Transport – shipping	Has corresponding existing measures in the SOM analysis (Yes/No)
9	Reduce re-suspension from sediments, by limiting restructuring of seabed to areas with low concentrations	64 (21) ○●●	Not assessed	Not assessed	No
10	Tighter allowed contamination levels during dredging activities	55 (16) ●●●	Not assessed	Not assessed	No
11	Perform dredging under conditions (low water pH and temperature) that lower desorption from sediments, i.e. Preferable during winter	30 (20) ○○●	Not assessed	Not assessed	No
19	Restrictions on anchoring zones in highly contaminated areas	Not assessed	Not assessed	31 (22) ○○●	No
20	In water hull cleaning regulation	Not assessed	Not assessed	61 (24) ○●●	No
21	Implement technologies for the degradation of TBT from sediment (e.g. chemical oxidation, bioremediation)	46 (20) ○●●	Not assessed	Not assessed	No
22	Boat washing restrictions	Not assessed	69.0 (33.1) ○●●	Not assessed	No
23	Monitoring of tin in leisure boat hulls	Not assessed	19.1 (11.7) ○○●	Not assessed	No
	Confidence	Moderate	Moderate	High	
	Number of experts	6	6	6	

Colour scale for the effectiveness of a measure type in percent (based on the expected value): 0-10%, 10-20%, 20-40%, 40-60%, 60-100%

Categories for the certainty of the effectiveness estimate (based on the relative size of the standard deviation to the expected value): low: ○○●, moderate: ○●●, high: ●●●

Data used: expert responses on effectiveness of measure types

Full activity names:

- Restructuring of seabed morphology (dredging, beach replenishment, sea-based deposit of dredged material)
- Tourism and leisure activities (boating, beach use, water sports, etc.)
- Transport – shipping (incl. anchoring, mooring)

Table 9.3 Effectiveness of measure types (%) in reducing the potential *input of PFOS.* The effectiveness of a measure type is the percent reduction in the pressure input resulting from a specific activity. The table depicts the expected effectiveness, and standard deviation is given in parenthesis.

Measure type ID	Activity Measure type	Transport – air	Urban uses	Industrial uses	Waste waters	Solid waste	Has corresponding existing measures in the SOM analysis (Yes/No)
24	Stockholm convention PFOS accepted use and specific exemptions limited to: insect baits, metal plating in a closed loop, fire-fighting foams	Not assessed	Not assessed	30 (22) ○○●	32 (29) ○○●	20 (25) ○○●	Yes
25	Stockholm convention PFOS lists no accepted uses or exemptions	Not assessed	Not assessed	54 (22) ○●●	33 (28) ○○●	29 (27) ○○●	Yes
26	Clean-up of contaminated sites	Not assessed	Not assessed	59 (29) ○●●	Not assessed	56 (27) ○●●	Yes
27	Restoration/upgrading of old landfill sites	Not assessed	Not assessed	Not assessed	Not assessed	50 (22) ○●●	Yes
28	Limits on PFOS concentrations in sludge used in commercial applications	Not assessed	Not assessed	Not assessed	39 (22) ○●●	Not assessed	No
29	Implement technologies to remove PFOS from wastewater (e.g. activated carbon or high-pressure membrane systems)	Not assessed	Not assessed	Not assessed	54 (22) ○●●	Not assessed	No
30	Stockholm convention ban on PFOS fire-fighting foams	Not assessed	68 (29) ○●●	Not assessed	Not assessed	Not assessed	No
31	Stockholm convention ban on PFOS in aviation hydraulic fluid	33 (29) ○○●	Not assessed	Not assessed	Not assessed	Not assessed	No*
	Confidence	Moderate	Moderate	Moderate	Moderate	Moderate	
	Number of experts	8	7	7	5-6	5-6	

*Implemented in the Baltic Sea region

Colour scale for the effectiveness of a measure type in percent (based on the expected value): 0-10%, 10-20%, 20-40%, 40-60%, 60-100%

Categories for the certainty of the effectiveness estimate (based on the relative size of the standard deviation to the expected value): low: ○○●, moderate: ○●●, high: ●●●

Data used: expert responses on effectiveness of measure types

Full activity names:

- Transport – air, including infrastructure
- Urban uses (land use), including storm water runoff
- Industrial uses (oil, gas, industrial plants)
- Waste waters (urban, industrial, and industrial animal farms; includes all waste streams entering wastewater systems e.g. microplastics, pharmaceuticals, etc.)

- Solid waste (e.g. land-based disposal of dredged material, landfill, solid waste streams)

Table 9.4 Effectiveness of measure types (%) in reducing the potential *input of diclofenac.* The effectiveness of a measure type is the percent reduction in the pressure input resulting from a specific activity. The table depicts the expected effectiveness, and standard deviation is given in parenthesis.

Measure type ID	Activity Measure type	Waste waters	Has corresponding existing measures in the SOM analysis (Yes/No)
32	Technical upgrade of wastewater treatment plants: e.g. granular activated carbon (GAC) adsorption, ozonation, UV light, nanofiltration etc	58 (22) ○●●	Yes
33	Improved application of existing WWT technologies (e.g. increasing the sludge retention time, use of both nitrification and denitrification treatment steps)	31 (28) ○○●	No
34	Improved pharmaceutical take-back schemes	33 (23) ○○●	No
35	Increase public awareness of pharmaceutical take-back schemes	35 (19) ○●●	No
36	Alter prescription practices to lower consumption (drug dosage, pack size, alternative medicine, convert OTC access to prescription)	50 (22) ○●●	No
	Confidence	High	
	Number of experts	7	

Colour scale for the effectiveness of a measure type in percent (based on the expected value): 0-10%, 10-20%, 20-40%, 40-60%, 60-100%

Categories for the certainty of the effectiveness estimate (based on the relative size of the standard deviation to the expected value): low: ○○●, moderate: ○●●, high: ●●●

Data used: expert responses on effectiveness of measure types

Full activity names:

- Waste waters (urban, industrial, and industrial animal farms; includes all waste streams entering wastewater systems e.g. microplastics, pharmaceuticals, etc.)

Which activities contribute to pressure inputs?

Table 10 shows the contribution of activities to the input of hazardous substances. Expert elicitation was used to estimate the activity-pressure input relationships for the pressure input concerning the input of mercury, PFOS, TBT and diclofenac, and the assessment was done at the level of the entire Baltic Sea.

19 different activities were identified to contribute to the input of the substances (see Table 10), with the number of activities varying between 5 to 12 (mercury 12, PFOS 9, TBT 10, and diclofenac 5). All the listed activities refer only to contributions originating within the Baltic countries. All contributions originating outside the Baltic countries, regardless of activity, are included under the category *activities and sources outside the Baltic Sea Region.* An additional category, *other/not determined*, captures remaining inputs not linked to a specific activity by experts. Due to variation in expert responses, contributions in this category could come from any of the 31 activities in the SOM analysis.

For mercury, the two activities that contribute the most to the pressure input are *activities and sources outside the Baltic Sea Region* and *non-renewable energy generation*. *Industrial uses* also have a considerable impact.

The activity *solid waste* contributes the most to the input of PFOS. Other important activities are *waste waters* and *activities and sources outside the Baltic Sea Region*. A significant portion of PFOS inputs come from undetermined sources (*other/not determined*).

For TBT, the activities with the highest contribution to the pressure input are *solid waste* and *restructuring of seabed morphology*, while the other activities have only minor contributions. A considerable amount of inputs come from undetermined sources (*other/undetermined*).

For diclofenac, the activity with the highest contribution to the pressure input is *waste waters* (51%). *Solid waste* (20%) is also an important activity. A significant part of the inputs come from undetermined sources (*other/not determined*).

The certainty of the activity-pressure input estimates is in most cases low and occasionally moderate, based on the standard deviations.

[Further discussion and interpretation of results, input from topic team]

Table 10. Activity-pressure input contributions (%). The activity-pressure input contributions show the percent share the activity contributes to the input of mercury, PFOS, TBT and diclofenac. Standard deviation is given in parenthesis.

Activity	Input of mercury	Input of PFOS	Input of TBT	Input of diclofenac
Agriculture			8 (13) ○○●	3 (4) ○○●
Forestry	1 (2) ○○●			
Non-renewable energy generation (fossil fuel and nuclear powerplants)	26 (18) ○○●			
Canalisation and other watercourse modifications (dams, culverting, trenching, weirs, large-scale water deviation)	1 (1) ○○●			
Transport – air, including infrastructure		2 (2) ○○●		
Urban uses (land use), including storm water runoff		10 (11) ○○●	4 (5) ○○●	5 (6) ○○●

Activity	Input of mercury	Input of PFOS	Input of TBT	Input of diclofenac
Industrial uses (oil, gas, industrial plants)	11 (15) ○○●	3 (4) ○○●		
Waste waters (urban, industrial, and industrial animal farms; includes all waste streams entering wastewater systems e.g. microplastics, pharmaceuticals, etc.)	1 (1) ○○●	18 (13) ○○●	4 (5) ○○●	51 (24) ○●●
Solid waste (e.g. land-based disposal of dredged material, landfill, solid waste streams)	6 (9) ○○●	28 (13) ○●●	24 (26) ○○●	20 (22) ○○●
Aquaculture – marine, including infrastructure	1 (1) ○○●			
Renewable energy generation (wind, wave and tidal power), including infrastructure			3 (5) ○○●	
Extraction of minerals (rock, metal ores, gravel, sand, shell)	5 (8) ○○●			
Extraction of oil and gas, including infrastructure (e.g. pipelines)	3 (4) ○○●	3 (4) ○○●		
Restructuring of seabed morphology (dredging, beach replenishment, sea-based deposit of dredged material)	6 (9) ○○●		23 (24) ○○●	
Tourism and leisure activities (boating, beach use, water sports, etc.)			4 (6) ○○●	2 (3) ○○●
Tourism and leisure infrastructure (piers, marinas)		2 (3) ○○●		
Transport – shipping (incl. anchoring, mooring)			4 (6) ○○●	
Transport – shipping infrastructure (harbours, ports, shipbuilding)	2 (4) ○○●	4 (6) ○○●	7 (9) ○○●	
Activities and sources outside the Baltic Sea Region	28 (21) ○○●	12 (10) ○○●	6 (9) ○○●	
Other/not determined	8 (9) ○○●	18 (8) ○●●	14 (15) ○○●	19 (23) ○○●
Number of experts	6	6	6	6

Colour scale for the contribution of the activity to the pressure input in percent (based on the expected value): 0-10%, 10-20%, 20-40%, 40-60%, 60-100%

Categories for the certainty of the activity-pressure input contribution estimate (based on the relative size of the standard deviation to the expected value): low: ○○●, moderate: ○●●, high: ●●●

Data used: expert responses on activity-pressure input contributions

What are the impacts of measure types?

The impacts of measure types show the impact of measure types on reducing the input of mercury, TBT, PFOS and diclofenac. They include the effectiveness of measure types and the contribution of activities to pressure input. Thus, the impact shows how much the measure type reduces the pressure input across all activities contributing to the pressure input and give indications on which measures could be the most relevant in addressing specific pressure inputs.

The measures types having the most impact on reducing the input of mercury *are local/state/national targets to eliminate coal fired energy production*, the *Minamata convention* and the *Paris agreement*.

Input of TBT is decreased most by *reducing re-suspension from sediments by limiting restructuring of seabed to areas with low concentrations*, *tighter allowed contamination levels during dredging activities*, and *implementing technologies for the degradation of TBT from sediment (e.g. chemical oxidation, bioremediation)*.

For the input of PFOS, *clean-up of contaminated sites*, the *Stockholm convention PFOS lists no accepted uses or exemptions,* and *restoration/upgrading of old landfill sites* are the most impactful measure types.

The input of diclofenac, the measure types having most impacts are *technical upgrade of wastewater treatment plants: e.g. granular activated carbon (GAC) adsorption, ozonation, UV light, nanofiltration etc.* and *altering prescription practices to lower consumption (drug dosage, pack size, alternative medicine, convert OTC access to prescription)*.

Detailed information on the impacts of measures are given in Annex 12.

[Discussion on existing measures and their impact, which existing measures are driving pressure input reductions, which HELCOM measures are important but not yet implemented]

Background of respondents

This section provides summarizing information on the number and background of experts contributing to the surveys to collect data for the analysis.

The number of experts contributing to the hazardous substance surveys is shown in Table 11, with the sub-topic division and geographic area presented in Table 12. For the effectiveness of measures survey for hazardous substances, altogether 14 survey responses with 18 contributing experts were received. One of the answers was a group response with five contributing experts. For the pressure-state survey, 18 responses from 20 experts were received, from all contracting parties except Russia. Two of the responses were group answers with two contributing experts in each. For the activity-pressure input survey, three responses were received with six contributing experts.

Table 11. Number of experts contributing to the hazardous substance surveys

Survey	DE	DK	EE	FI	LT	LV	PL	RU	SE	Total
Effectiveness of measures	-	1	3	6	1	-	2	-	5	18
Pressure-state linkages	1	1	3	5	1	3	1	-	5	20
Activity-pressure input contributions	-	1	1	4	-	-	-	-	-	6

Table 12. Number of responses to the hazardous substance surveys

Survey	Sub-topic	Geographic area	Response count
Effectiveness of measures	Mercury	Whole Baltic	11
	TBT	Whole Baltic	12
	PFOS	Whole Baltic	10
	Diclofenac	Whole Baltic	12
Pressure-state linkages	Mercury	Whole Baltic	12
	TBT	Whole Baltic	8
	PFOS	Whole Baltic	8
	Diclofenac	Whole Baltic	12

For the experts participating in the effectiveness of measures and the pressure state survey, more detailed information about their background is available (Table 13). Experts stated most often environmental research, chemistry, contamination and marine science as their respective field. About half of the experts had 10-20 years of experience in their field, while only 5-6% had 0-2 years of experience. Experts represented research institutions, environmental companies, government institutes, or ministries.

Table 13. Years of experience in the field for the litter effectiveness of measures survey

	Effectiveness of measures		Pressure-state	
Years	**Number of experts**	**Share of experts**	**Number of experts**	**Share of experts**
0-2 years	1	6 %	1	5 %
3-5 years	2	11 %	2	10 %
5-10 years	3	17 %	3	15 %
10-20 years	9	50 %	11	55 %
over 20 years	3	17 %	3	15 %

Discussion

Impact of alternative scenarios for development of human activities

The detailed results are presented for the most likely development scenario for the extent of human activities in 2016–2030. In addition, three other development scenarios were estimated: no change, low change and high change scenarios. These scenarios cover 9 out of the 31 activities in the SOM analysis. The extent of other activities is assumed to remain constant in all scenarios.

As activities contribute to pressure inputs, their assumed change over time affects the pressure input reductions and probability to achieve GES or state improvements. The impact depends on to what extent the activities contributing to the specific pressure input are covered in the change scenarios. For hazardous substances, the coverage of activities that contribute to pressure inputs in the change scenarios ranges from very low to high, depending on the substance.

Overall, the impact of alternative development scenarios is limited in the case of the four substances. For mercury, the development scenario has only negligible impacts on pressure input reductions and probability to achieve GES, as almost all activities contributing to the input of mercury are excluded from the 9 activities that have development scenarios, and are thus assumed to remain constant in all scenarios.

For TBT, the only impact on inputs comes from the changes in human activities, and therefore the scenario has a clear effect on pressure input reductions, despite the fact that most activities contributing to TBT are also expected to remain constant. When activities are assumed to stay constant, there are no projected changes in the input of TBT. When a high change scenario is assumed, the input of pressures increases compared to the most likely scenario. However, this change in pressure inputs does not affect the probability to achieve GES.

For PFOS, the change scenarios have a negligible influence on the input reductions, because the activities contributing to the input of PFOS are not projected to change significantly in any of the scenarios.

Similarly, for diclofenac, the impact of the assumption on future development of activities is limited due to minor changes in the extent of human activities in all scenarios, although the change scenarios cover over half of the activities contributing to its input.

Impact of using literature data on effectiveness of measures

In addition to survey data from experts, literature data on the effectiveness of measures has been compiled. The literature data points have been used in a similar way as the expert survey responses, and when it has been available, it has been used to replace the expert estimates of the effectiveness of the measure type. However, literature estimates are not available for all measure types. Thus, it is not possible to implement the model estimation and provide the results relying entirely on the literature data on effectiveness of measure types. Thus, the model including the literature estimates is a combination of literature and expert data on effectiveness of measure types. The origin of other data components is not affected.

For hazardous substances, 72 estimates from 23 studies could be included in the SOM model. The projected pressure input reductions from existing measures are not significantly affected by the inclusion of literature data for any of the substances. Thus, the results on sufficiency of measures to achieve GES do not change.

Evaluation of quality and confidence

The SOM analysis for hazardous substances in complete in the sense that it has been possible to evaluate the sufficiency of existing measures to achieve GES. All elements of results have been presented in a quantitative format, as the data have been deemed to suffice for that. It is worth noting that the results are not representative for all hazardous substances, but only to the four substances covered in the analysis.

The overall certainty of the assessment for hazardous substances could generally be characterized as moderate. The number of expert responses is relatively high, and experts from eight coastal countries have contributed to some part of the assessment. For the individual results, average certainty is low for the activity-pressure input contributions, moderate for effectiveness of measures types, and low-moderate to projected reductions in pressure inputs. There is also considerable uncertainty about the required pressure reductions to achieve GES. These uncertainties should be kept in mind, in particular when examining the numeric estimates. Quality and precision could potentially be improved with the collection of additional expert responses.

The most common confidence level experts reported for their own evaluations are moderate for effectiveness of measures, high for significance of pressures to state components, and moderate for required pressure reductions. This indicates relatively confidence in the survey responses.

Further, future development scenarios have been constructed only for few activities contributing to the input of the substances, and several important activities have been assumed to remain constant until 2030. This applies in particular to mercury. The relevance of the analysis could be improved by covering additional activities in the change scenarios.

There were some technical challenges that affected the survey implementation. Firstly, there was a problem in the survey software for the effectiveness of measure types survey that resulted in losing some responses. The original responses became often unusable, as it was not possible to identify which items had been skipped on purpose and which were lost data. This issue was addressed by sending follow-up invitations for experts to review and, when needed, complement their original saved response. Not all experts participated in the review, and thus their response had to be deleted from the final sample. Secondly, the simultaneous assessment of effectiveness of a measure type and certainty of that effectiveness proved in some cases difficult, as it required placing non-quantitative dots in a coordinate system to generate quantitative estimates. The dots were translated into effectiveness and certainty values between 0 and 100. Some experts would have preferred that the quantitative estimates would have been visible and could have been transparently influenced.

When interpreting the results, the assumptions and generalizations that were made when collecting the input data and defining and using the data on activity-pressure input contributions, measure type effectiveness and pressure-state linkages need to be taken into account. The input data are based mainly on expert elicitations rather than existing models and data, and reflect substantial uncertainty. For more information on the SOM methodology, data collection and assumptions, see this document.

[Add information on the credibility of the results and how they should (and should not) be interpreted, input from topic team]

Reflection on measure types

Generally, the measure types for hazardous substances appear to have struck a balance between specificity and general applicability, and as a whole do not show any systemic design flaws. However, for individual measure types, potential improvements can still be identified.

For mercury, measure types 1 and 2, covering mercury from non-renewable energy generation, very likely overlap in that measure type 1 can be implemented under the actions of measure type 2. The survey question for measure type 1 may also be poorly suited to elicit an informed response from experts. Estimated effectiveness could include any remaining mercury emissions from non-renewable energy generation (heating oil, waste combustion, peat, etc.) and the likelihood of the target being met. Whether this additional variability is desirable is unclear, but the measure type should be reviewed in any future work. Measure types 12 and 13 cover decreased use of fossil fuels in the transportation industry, but their formation leaves the quantity of reduction unclear which undoubtably increases the uncertainty of any estimate. This has been a point of difficulty when designing the measure types generally and methods of more precisely addressing this will be important moving forward. Measure type 18 has value as an indication that experts see potential improvements in further restrictions on mercury disposal, but fails to provide any context for the experts or the results as to what those restrictions might be. This reduces its applicability when used in the model. Generally, these types of catch-all measure types are likely to remain but their prominence should be reduced as much as possible by providing an otherwise comprehensive list of more definitive measure types.

For PFOS, measure types 26 and 27 concern local actions to reduce release from contaminated sites and it is unclear how these estimates should be interpreted. These measure types are part of a broader set of measure types addressing localized action where the reported estimates are difficult to link to existing measures (e.g. measure types 1, 4, 21, 29, 32, 33). Without an understanding of the significance of any particular local action to the broader input, use of these measure types in the model is not ideal. However, they do provide a clear indication of the importance of these measure types to the overall reduction of inputs. This spatial aspect of local measures must be a priority for future effort within this framework.

Finally, measure types 9, 10 and 11 concerning resuspension of substances from dredging activities have been applied across several substances. This is not a concern in and of itself, but should be a focus for review prior to use in any future work.

[Add further reflection on the measure types, input from topic team]

Lessons learned

Many different approaches have been applied to specific topics within the general SOM framework and hazardous substances is no exception. Some of these approaches have worked well while others have not.

The choice to focus on specific substances rather than categories of pollutants removed any chance of broad coverage of the topic of hazardous substances. However, this trade-off has clearly been the correct choice, as any attempt at broad coverage would have been completely infeasible due to the structure of existing measures, difficulty in identifying an appropriate expert pool, and poor applicability of any model outcome.

Hazardous substances took a unique approach to informing experts of the current state of substance concentration in the Baltic Sea. Instead of a one-out-all-out approach (used in e.g. coastal fish assessment), conducting separate assessments for different spatial areas (used in all topics except NIS and hazardous substances) or providing a Baltic wide assessment (not consistent with the HELCOM HOLAS II assessment structure and rejected by topic experts as inappropriate), experts were provided with a link to the latest indicator reports containing detailed reviews of the latest assessments for each substance. This approach performed more poorly than the options used for other topics and in the future either a one-out-all-out approach and/or multiple spatial assessments should be used.

The pressure list for the SOM analysis was not adapted for any topics as the list of 23 pressures was considered comprehensive enough to cover all topics. However, this caused difficulties for hazardous substances where assumptions had to be made to allow for a complete analysis (see the Topic specific model structure, assumptions and challenges section). This issue could be solved by amending the pressure list for hazardous substances to include substance specific pressures (e.g. mercury pollution) rather than only the more general pollution pressures (e.g. heavy metal pollution).

Future work should also bring more clarity to the division of the input of a pollutant and its environmental concentration. While the pressure inputs are clearly delineated between the two (e.g. input of mercury), the hazardous substance pressures (e.g. heavy metal pollution) are not. Further it is not clear the most appropriate formulation for the pressures given the structure of the SOM analysis (input and current concentrations versus just input). Further work on this aspect is needed.

Use of results, implications and future perspectives

[Add information on how and to what purposes the results could be used and on the practical implications of the work, input from topic team]

[Include a short description about the issue of long-range emissions, input from topic team]

[Add future perspectives and what are the information gaps to be tackled in the future, input from topic team]

It is worth noting that the results are not representative for all hazardous substances, but only to the four substances covered in the analysis. Thus, the results should not be generalized to broader sets of hazardous substances.

Annexes

Annexes 1–9 contain the expert surveys as well as information on the measure types and the literature review. They are available on the SOM Platform workspace.

Annexes 10–12 contain graphs and tables that provide additional information and perspectives on the results.

Annex 1 Activity-pressure input survey template

Excel used as a template for receiving data for the activity-pressure input survey.

Annex 2 Modified activity list (if modified)

The topic uses the standard activity list, so no modified activity list is available.

Annex 3 Measure types list

PDF containing the measure types used in the assessment of the effectiveness of measures for *Hazardous substances*. Document includes examples of existing measures that if implemented would be included in the corresponding measure type.

Annex 4 Linking existing measures to measure types

Excel containing the identified existing measures and their relationship to the measure types used in the SOM analysis.

Annex 5 Literature review search terms

Excel containing the search terms used during the literature review on effectiveness of measures for *Hazardous substances*.

Annex 6 Literature review summary

Excel document containing the effectiveness of measures data retrieved from the literature review.

Annex 7 Topic structure

Excel containing the relationships between measure types, activities, pressure inputs, state components, and sub-basins. Also contains information on GES thresholds.

Annex 8 Effectiveness of measures survey

PDF of the Effectiveness of measures survey for *Hazardous substances*.

Annex 9 Pressure-state survey

PDF of the Pressure-state survey for *Hazardous substances*.

Annex 10 Supplementary results for required pressure reductions

This annex presents the probability density functions of required pressure reductions to achieve GES based on responses to the expert survey questions. The graph shows the probability distribution of the pooled expert responses on how much pressures should be reduced to achieve GES. Pressure reduction is presented on the x-axis (0-100%) and probability on the y-axis. The probability density function presents the probability of the pressure reduction falling within a particular range of values. This probability is given by the integral of the probability density over that range—that is, it is given by the area under the density function but above the horizontal axis and between the lowest and greatest values of the range.

The graphs have multiple peaks and the distributions are wide, which indicate that expert have varying views on the pressure reductions required to achieve good state. For example, some experts are of the opinion that reductions around 10% would be sufficient to achieve GES for PFOS, while others consider that reductions around 70-80% are required.

[updated graphs to be included later]

[improve the outlook of the graphs, include in all figures and graphs the number of experts contributing to the result, include standard deviations or confidence intervals in the graphs, where appropriate]

Annex 11 Supplementary results for effectiveness of measures

[updated graphs to be included later]

[improve the outlook of the graphs, include in all figures and graphs the number of experts contributing to the result, include standard deviations or confidence intervals in the graphs, where appropriate]

Annex 12 Impacts of measure types

Table A1. Impacts of measure types (%) in reducing the input of mercury, TBT, PFOS and diclofenac. The impact shows how much the measure type reduces the pressure input across all activities contributing to the pressure input.

Substance	Measure type	Impact of measure type (%) Mean (standard deviation)
Mercury	Local/state/national targets to eliminate coal fired energy production	18 (16)
	Minamata convention	14 (12)
	Paris agreement	13 (12)
	Continued application of the EU Emissions Trading System	9 (9)
	Updated EU mercury storage regulation	3 (6)
	Sand capping or removal of existing cellulose deposits	3 (6)
	Reduce re-suspension from sediments, by limiting restructuring of seabed to areas with low concentrations	3 (6)
	Ban on manufacturing processes where mercury is used as a electrode (e.g. in chlor-alkali production; Na or K methylate/ethylate production	3 (6)
	Ban on manufacturing processes where mercury or mercury compounds are used as a catalyst (e.g. In vinyl chloride monomer production	3 (6)
	Further restrictions on storage and disposal of waste/dredged material containing mercury	3 (5)
	Tighter allowed contamination levels during dredging activities	3 (4)
	Ban on export of Mercury (II) sulphate ($HgSO_4$, CAS RN 7783-35-9) and Mercury (II) nitrate ($Hg(NO_3)_2$, CAS RN 10045-94-0)	3 (6)
	5% increase in EU mandatory WEEE recycling levels	2 (4)
	Eurasian Economic Union restrictions on hazardous substances in electrical products	2 (4)
	Perform dredging under conditions (low water pH and temperature) that lower desorption from sediments, i.e. Preferable during winter	2 (3)
	EU mandatory use of dental amalgam separators retaining at least 95% of amalgam particles	0 (0)
TBT	Reduce re-suspension from sediments, by limiting restructuring of seabed to areas with low concentrations	15 (17)
	Tighter allowed contamination levels during dredging activities	13 (14)
	Implement technologies for the degradation of TBT from sediment (e.g. chemical oxidation, bioremediation)	11 (12)
	Perform dredging under conditions (low water pH and temperature) that lower desorption from sediments, i.e. Preferable during winter	7 (9)
	In water hull cleaning regulation	2 (4)
	Boat washing restrictions	2 (4)
	Restrictions on anchoring zones in highly contaminated areas	1 (2)
	Monitoring of tin in leisure boat hulls	1 (1)
PFOS	Clean-up of contaminated sites	18 (11)
	Stockholm convention PFOS lists no accepted uses or exemptions	16 (11)
	Restoration/upgrading of old landfill sites	14 (9)
	Stockholm convention PFOS accepted use and specific exemptions limited to: insect baits, metal plating in a closed loop, fire-fighting foams	12 (11)

Substance	Measure type	Impact of measure type (%) Mean (standard deviation)
	Implement technologies to remove PFOS from wastewater (e.g. activated carbon or high-pressure membrane systems)	10 (8)
	Limits on PFOS concentrations in sludge used in commercial applications	7 (7)
	Stockholm convention ban on PFOS fire-fighting foams	7 (8)
	Stockholm convention ban on PFOS in aviation hydraulic fluid	1 (1)
Diclofenac	Technical upgrade of wastewater treatment plants: e.g. granular activated carbon (GAC) adsorption, ozonation, UV light, nanofiltration etc	30 (20)
	Alter prescription practices to lower consumption (drug dosage, pack size, alternative medicine, convert OTC access to prescription)	26 (17)
	Increase public awareness of pharmaceutical take-back schemes	18 (13)
	Improved pharmaceutical take-back schemes	17 (14)
	Improved application of existing WWT technologies (e.g. increasing the sludge retention time, use of both nitrification and denitrification treatment steps)	16 (17)

Data used: activity-pressure input contributions, effectiveness of measure types

References

Blom, C. and Hansen, L. (2015). Analysis of per- and polyfluorinated substances in articles. Nordic Working Papers. 2015:911HELCOM (2018a): State of the Baltic Sea – Second HELCOM holistic assessment 2011-2016. Baltic Sea Environment Proceedings 155.

HELCOM (2020). Methodology for the sufficiency of measure analysis. Updated 28 October 2020. Available at: https://portal.helcom.fi/workspaces/ACTION-164/Public%20documents/Methodology_for_the_SOM_analysis.pdf

HELCOM (2018b) Metals (lead, cadmium and mercury). HELCOM core indicator report. Online. 30.6.2020, https://helcom.fi/wp-content/uploads/2019/08/Metals-HELCOM-core-indicator-2018.pdf

HELCOM (2018c) Perfluorooctane sulphonate (PFOS). HELCOM core indicator report. Online. 30.6.2020, https://helcom.fi/wp-content/uploads/2019/08/Perfluorooctane-sulphonate-PFOS-HELCOM-core-indicator-2018.pdf

HELCOM (2018d) Tributyltin TBT and imposex. HELCOM core indicator report. Online. 30.6.2020, https://helcom.fi/wp-content/uploads/2019/08/Tributyltin-TBT-and-imposex-HELCOM-core-indicator-2018.pdf

HELCOM (2018e) Diclofenac. HELCOM pre-core indicator report. Online. 30.6.2020, https://helcom.fi/wp-content/uploads/2019/08/Diclofenac-HELCOM-pre-core-indicator-2018.pdf

HELCOM (2018f). HELCOM Thematic assessment of hazardous substances 2011-2016. Supplementary report to the HELCOM 'State of the Baltic Sea' report.

HELCOM, (2018g). Inputs of hazardous substances to the Baltic Sea. Baltic Sea Environment Proceedings No. 161

HELCOM (2018h). Polyaromatic hydrocarbons (PAHs) and their metabolites. HELCOM core indicator report

HELCOM 2018i. Radioactive substances: Cesium-137 in fish and surface seawater. HELCOM core indicator report

Johansson, J. and Undeman, E. 2020. Perfluorooctane sulfonate (PFOS) and other perfluorinated alkyl substances (PFASs) in the Baltic Sea – Sources, transport routes and trends. Helcom Baltic Sea Environment Proceedings n°173

McLachlan, M. & Undeman, E. 2020. Dioxins and PCBs in the Baltic Sea. Helcom Baltic Sea Environment Proceedings n°171

Sobek, A., Bejgarn, S., Ruden, C., & Breitholtz, M. (2016). The dilemma in prioritizing chemicals for environmental analysis: known versus unknown hazards. Environmental Science: Processes & Impacts, 18(8), 1042-1049.

Undeman, E. 2020. Diclofenac in the Baltic Sea – Sources, transport routes and trends. Helcom Baltic Sea Environment Proceedings n°170

Undeman, E. and Johansson, J. 2020. Polybrominated diphenyl ethers (PBDEs) in the Baltic Sea – Sources, transport routes and trends. Helcom Baltic Sea Environment Proceedings n°172

UNESCO and HELCOM. 2017. Pharmaceuticals in the aquatic environment of the Baltic Sea region – A status report. UNESCO Emerging Pollutants in Water Series – No. 1, UNESCO Publishing, Paris.

WHO 2008. Guidance for identifying populations at risk from mercury exposure. UNEP DTIE Chemicals Branch and WHO Department of Food Safety, Zoonoses and Foodborne Diseases, Geneva, Switzerland