Document title	Use of survey results from expert elicitation in the SOM model
Code	2-3
Category	CMNT
Agenda Item	2 - Approach to analyse sufficiency of measures
Submission date	6.9.2019
Submitted by	ACTION WP6, Secretariat
Reference	

Background

Document 2-2 outlines the approach and steps for analysing sufficiency of measures, including the use of surveys to collate expert-based estimates. This document presents in detail how the results of the survey will be treated and used in the SOM model. This includes: identifying main pathways for pressures using activity-pressure-linkages (step 3), estimation of effects of measures (step 4), calculation of the total effect of measures on a pressure (step 4), linking reduced pressures with state components (step 6), and assessing sufficiency of measures (step 7), including the probability for reaching good environmental status.

Action requested

The Meeting is invited to:

- take note of how survey results will be treated and used in the SOM model,
- take note of how the probability for reaching good environmental status is calculated,
- guide the further development of the analyses as necessary.

Use of survey results from expert elicitation in the SOM model

1. Identifying main pathways for pressures (step 3)

Using expert evaluation to identify main pathways for pressures (activity-pressure contribution surveys) three-point estimates (min-%, max-%, most likely-%) are provided by each expert for each sub-basin group (these depend on the topic and can consist of multiple HELCOM scale 2 sub-basins) that they are willing to estimate based on their expertise/knowledge. For some topics this information can be based on existing empirical data or existing empirical data can be used to complement the expert surveys.

The base case results are calculated using activity-pressure contribution probability distributions that are defined based on the surveyed three-point estimates. This allows the comparison of results among different topics and also the consistent assessment of pressure reductions for different state components that are affected by multiple pressures. The distributions are aggregated from the three-point estimates of individual experts, and several alternative assumptions of the shape of the aggregated distributions can be made (e.g. triangular or beta-distribution). Sensitivity analyses can later be made deviating from the base case assumptions, by using alternative distribution types that are more representative for different topics than those based on three-point estimates and that can differ among topics, if such distributions based on empirical data are available. The same principle applies for the other distributions used in this analysis.

The aggregated probability distribution of an activity ($\boldsymbol{j}$) - pressure ($\boldsymbol{i}$) contribution $\boldsymbol{C}$ for a basin ($\boldsymbol{k}$) is

$$\boldsymbol{fc_{i,j,k}(C)}$$

2. Estimation of effects of measures (step 4)

The effectiveness of measures is surveyed by measure type for each activity that is contributing significantly to a pressure (as determined in step 3). The effectiveness is not assessed individually for actual measures, but for more aggregated measure types which are defined based on existing measures. This is done for several reasons: i) there are too many measures to assess them individually, ii) the available information on existing measures is incomplete and asymmetric, which could jeopardize equal assessment of measure effectiveness among topics, different countries and policy schemes, and iii) the surveyed measure type effectiveness can be applied to assess the effectiveness of new measures.

The effectiveness of measures survey consists of two parts. The first part is a grid question where different measure types are located based on their (relative) effectiveness (x-axis: no effect-highest effect) and certainty of their effectiveness (y-axis: uncertain-certain) to reduce a given pressure from a specific activity. Uncertainty here refers to the objective uncertainty arising from the level of scientific evidence on measure effectiveness and also on the variation of measures that comprise one measure type. The grid-question is asked for each significant activity contributing to a certain pressure, based on the activity-pressure contributions (for an example of a grid question on the effectiveness of measure types targeting one activity-pressure combination see Figure 1). Here the set of measure types related to each pressure $\boldsymbol{N_i}$ includes all possible measure types affecting that pressure, but some of these might not be relevant for all activities. Thus, the measure types included in the question may be subsets of $\boldsymbol{N_i}$. It is assumed that the measure type effectiveness to reduce a given pressure from a certain activity is the same for the whole Baltic Sea. However, one has to remember that activity-pressure contributions and actual pressure levels can vary spatially, which means that the absolute or even the relative effects of certain measure type on the total pressure reduction likely differ between the basins. This is demonstrated later in this document where the total effect of measures on a pressure is assessed for one basin.

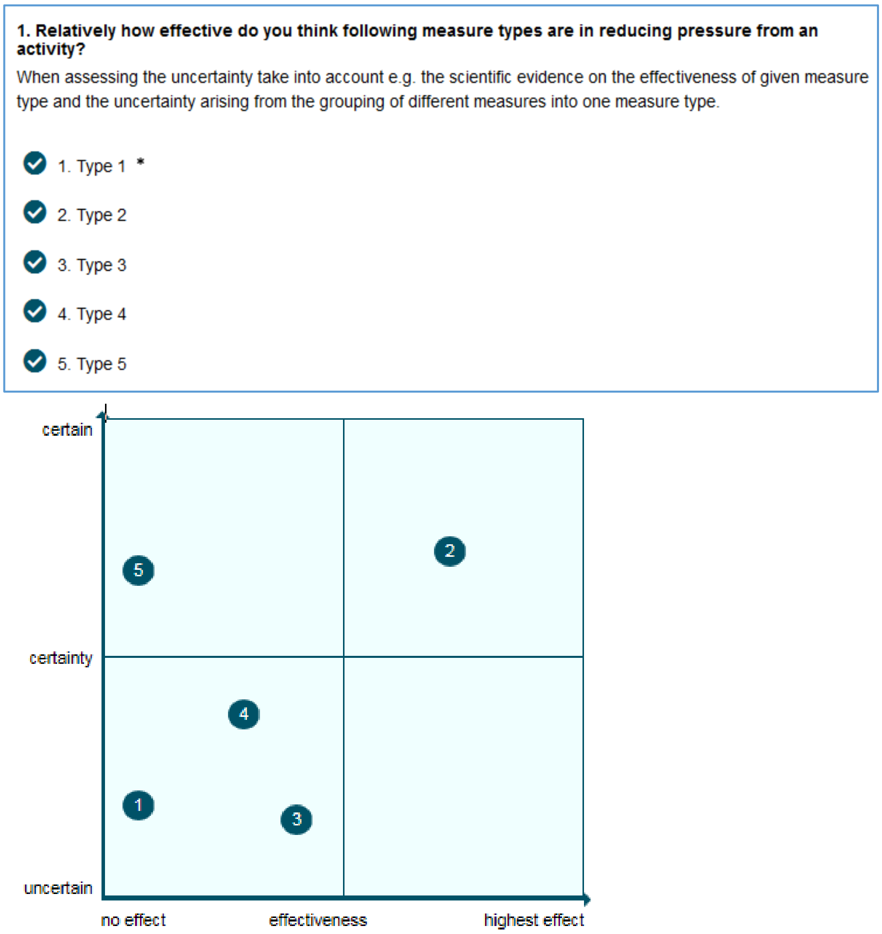


Figure 1. Grid-question on the (relative) effectiveness of measure types and the certainty of the effectiveness

The relative effects of different measure types with respect to the most effective measure type (the one the most right on the x-axis) are used to scale the measure type effects. They are defined for each measure type of each grid-question by

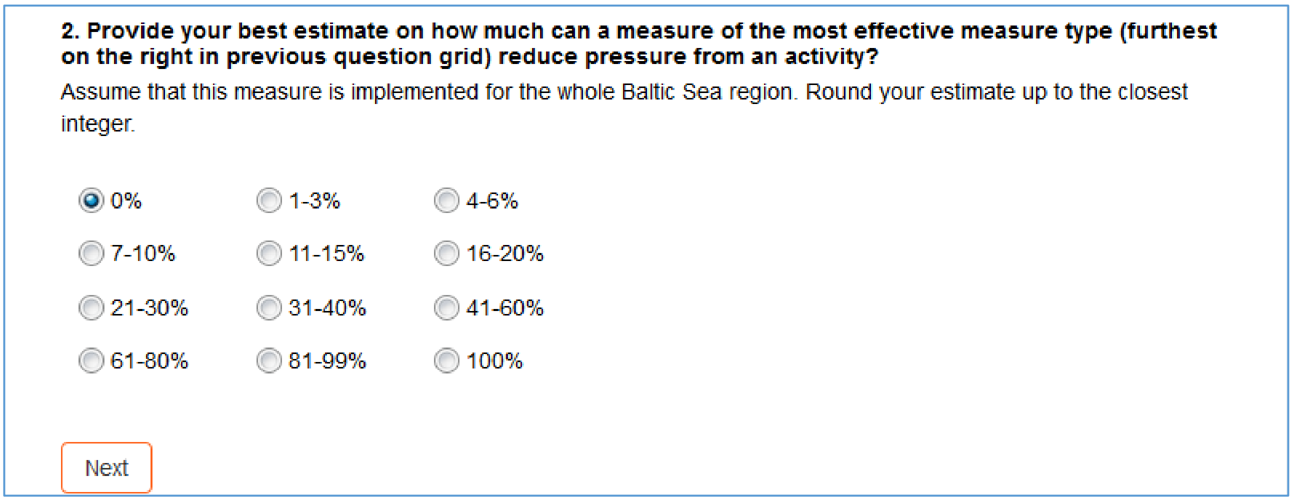
$$E_{i,j,n} = \frac{x_{i,j,n}}{x_{i,j,max}} \qquad (1)$$

where $x_{i,j,n}$ is the position on the x-axis and max refers to the most effective measure type (most right on the x-axis).

The uncertainty values (position on the y-axis) are used to define a range of the effectiveness of the measure type in reducing the pressure from an activity. Minimum certainty (uncertain) is assumed to mean that all possible effectiveness levels from no effect to highest effect are possible and the most likely effectiveness is the place on the x-axis where an expert has placed the relative effectiveness value with respect to the other measures. Maximum certainty (certain) is assumed to mean that the effectiveness of measure type always equals the most likely value, and thus there is no range but only a point value of effectiveness. There are alternative ways to treat cases where the uncertainty is high and will be addressed in the next development phase. The relative minimum and maximum effects that different measure types can take with respect to the most effective measure type can be calculated in the same way as in formula

(1): $E_{i,j,n,L} = \frac{x_{i,j,n,L}}{x_{i,j,max}}$ and $E_{i,j,n,H} = \frac{x_{i,j,n,H}}{x_{i,j,max}}$, where $\boldsymbol{x_{i,j,n,L}}$ and $\boldsymbol{x_{i,j,n,H}}$ are the lowest and highest end of the effectiveness value range for the measure type respectively.

The second part of the expert survey related to measure effectiveness (Figure 2) asks, in percentages, how much the most effective measure type (most right on the x-axis) reduces the pressure from the activity. The most likely effect of the most effective measure type can be defined as the mean of the given percentage reduction range or as a distribution of the values in that range. Using the mean, we can denote this effect by $\bar{R}_{i,j,max}$. The most likely effect of other measure types can be estimated as a product of the expected effect of the most effective measure type and the relative effect of a measure type with respect to the most effective measure type $\bar{R}_{i,j,n} = E_{i,j,n} \times \bar{R}_{i,j,max}$. The minimum and maximum effects for different measure types are calculated in a similar fashion but using $E_{i,j,n,L}$ and $E_{i,j,n,H}$ respectively. These effect ranges (most likely, minimum and maximum effects) define three-point estimates for each survey response. Again, the probability distributions for measure type effects as %-reduction in pressures from activities are aggregated from the three-point estimates of individual experts, and several alternative assumptions of the shape of the aggregated distributions can be made (e.g. triangular and beta-distribution). The probability distribution of a %-pressure reduction effect $\boldsymbol{R}$ of a measure type $\boldsymbol{n}$ on a pressure i from an activity j is $\boldsymbol{r_{i,j,n}(R)}$. Again, scientific evidence in literature on the measure type effects can also be used to define alternative distributions for measure type effects, if such data are available.



2. Provide your best estimate on how much can a measure of the most effective measure type (furthest on the right in previous question grid) reduce pressure from an activity?

Assume that this measure is implemented for the whole Baltic Sea region. Round your estimate up to the closest integer.

0%	1-3%	4-6%
7-10%	11-15%	16-20%
21-30%	31-40%	41-60%
61-80%	81-99%	100%

Next

Figure 2. How much can the most effective measure type reduce pressure from an activity?

3. Calculating the total effect of measures on a pressure (step 4)

The total pressure reduction effect $\boldsymbol{T_{k,i}}$ of measures on a pressure $\boldsymbol{i}$ in a basin $\boldsymbol{k}$ is calculated as a sum of all effects of measures affecting pressure $\boldsymbol{i}$ in basin $\boldsymbol{k}$ multiplied by their respective activity pressure contributions

$$\boldsymbol{T_{k,i}} = \sum_{j \in A_{k,i}} \boldsymbol{C_{i,j,k}} \sum_{n \in N_i} \sum_{m=1}^{M_{k,j,i,n}} \boldsymbol{R_{i,j,n}}$$

where $\boldsymbol{C_{i,j,k}}$ is the contribution of an activity j on pressure i in basin k, $\boldsymbol{A_{k,i}}$ is the set of significant activities causing pressure i in basin k, $\boldsymbol{N_i}$ is the set of all measure types linked to pressure i, and $M_{k,j,i,n}$ denotes the number of measures of measure type n affecting pressure i from activity j in basin k, and $\boldsymbol{R_{i,j,n}}$ is the pressure reduction effect of the given measure type n on pressure i from activity j. Measure effects of individual measures are thus defined by the effectiveness of the measure type that they belong to.

The distribution of the total pressure reduction (in %) for pressure i in basin k is defined by drawing a large number (likely N=100 000) of random values for $\boldsymbol{C_{i,j,k}}$ and $\boldsymbol{R_{i,j,n}}$ from their respective activity-pressure contribution probability distribution $\boldsymbol{fc_{i,j,k}(C)}$ and measure type effect probability distribution $\boldsymbol{r_{i,j,n}(R)}$ [1], and then for each 100 000 sets of values[2] drawn from these distributions calculating the total reduction in pressure $\boldsymbol{T_{k,i}}$.

If a measure affects only certain parts of some basin (for example national measures) then the effect is multiplied by the area of that part of the basin divided by the area of the whole basin. A probability distribution is defined for each pressure reduction in % based on the N=100 000 calculated pressure reduction effects. These distributions take into account the uncertainty in the activity-pressure contributions, as well as in the effectiveness of the measure types. These distributions allow for calculating the expected pressure reductions, constructing confidence intervals for pressure reductions, and calculating the probability to reach a specific pressure reduction.

4. Linking reduced pressures with state components (step 6)

The first survey question on the pressure-state linkage asks the experts to identify the most significant pressures preventing the state variable (such as the abundance of some species) from reaching good environmental state (Figure 3). These are asked separately for each assessed basin. The pressures are weighted based on their proportion of the total significance of pressures:

$$W_{i,k,s} = \frac{\boldsymbol{y_{i,k,s}}}{\sum y_{k,s}}$$

where $\boldsymbol{y_{i,k,s}}$ is the significance score (0-5, 0 being not significant in Figure 3) of the given pressure $\boldsymbol{i}$ in basin $\boldsymbol{k}$ for state variable $\boldsymbol{s}$ and $\sum y_{k,s}$ is the sum of all significance scores of all significant pressures for state $\boldsymbol{s}$ in basin $\boldsymbol{k}$. The averages of elicited pressure significance scores are likely used to define $\boldsymbol{y_{i,k,s}}$.

[1] The values of $\boldsymbol{R_{i,j,n}}$ can also be drawn independently for each measure belonging to a certain measure type.

[2] The drawn values for activity-pressure contributions $\boldsymbol{C_{k,j,i}}$ are normalized for each set of values, so that the sum of contributions of different activities to a pressure equal 100%. The drawn values of pressure reduction $\boldsymbol{R_{i,j,n}}$ may in theory exceed 100%, but in such cases the pressure reduction is assumed and set to be 100%.

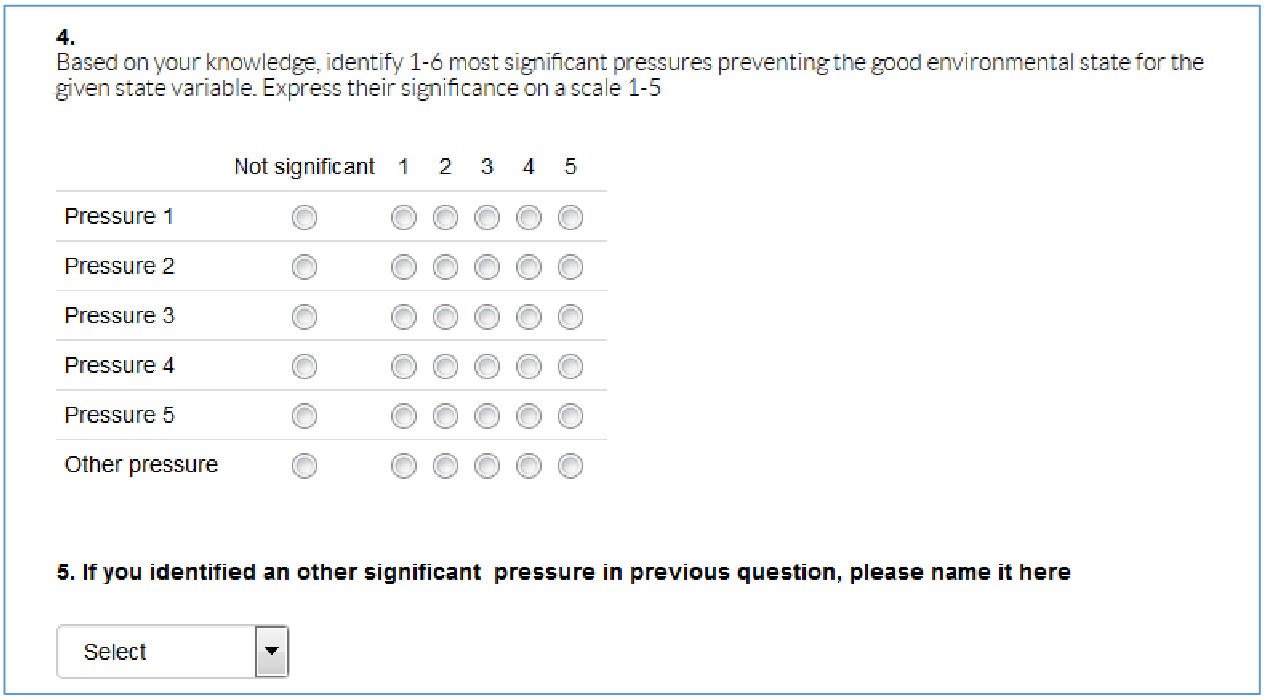


Figure 3. Question to identify the most important pressures preventing the good state for a state variable

The second survey question about pressure-state linkage (Figure 4) asks how much the most significant pressure needs to be reduced in order to reach good state for the state variable, assuming that other pressures are reduced the same percentage amount. Here, it is also assumed that there is no time lag between the reduction in pressure and the improvement in state. If the good state or/and current state can be quantified, these values are used when phrasing the questions for the pressure-state linkage. An alternative version of the question will be developed for those topics for which there are no agreed GES values. The questions could be rephrased, for example, to ask for the pressure reduction required to reach a specific state improvement level instead of GES.

The second question about the pressure-state linkage is again asked as a value range (most likely, minimum, maximum), where three-point estimates are provided by each expert. From these values, a cumulative distribution function can be defined that represents the probability of reaching a good state for different % reductions in total pressure (all significant pressures reduced by the same % amount) $\boldsymbol{FS_{k,s}(TPR)}$, where $\boldsymbol{TPR}$ is the reduction in total pressure. In reality it is very unlikely that all pressures are reduced by the same proportion, and thus the reduction in total pressure $\boldsymbol{\widetilde{TPR}_{k,s}}$ can be approximated using the pressure weights based on the significance scores from the question in Figure 3.

$$\boldsymbol{\widetilde{TPR}_{k,s}} = \sum_{\boldsymbol{i \in I_s}} W_{i,k,s} \boldsymbol{T_{k,i}}$$

where $\boldsymbol{i} \in \boldsymbol{I_s}$ is the set of significant pressures for state $\boldsymbol{s}$. By plugging the approximated reduction in total pressure into the function of reaching a good state for different % reductions one can estimate the probability of reaching a good state for state variable $\boldsymbol{s}$. If an expected value of total pressure reduction is applied to study how reductions in pressures increase the probability to reach good state, then cumulative distribution function $\boldsymbol{FS_{k,s}}$ can be used to define the expected probability to reach good environmental state. Whereas, if total pressure reduction is defined as a distribution, then probability distribution of reaching good state with a probability can be assessed and from that it is possible to estimate what is the likelihood that the probability of reaching a good state is at least X%.

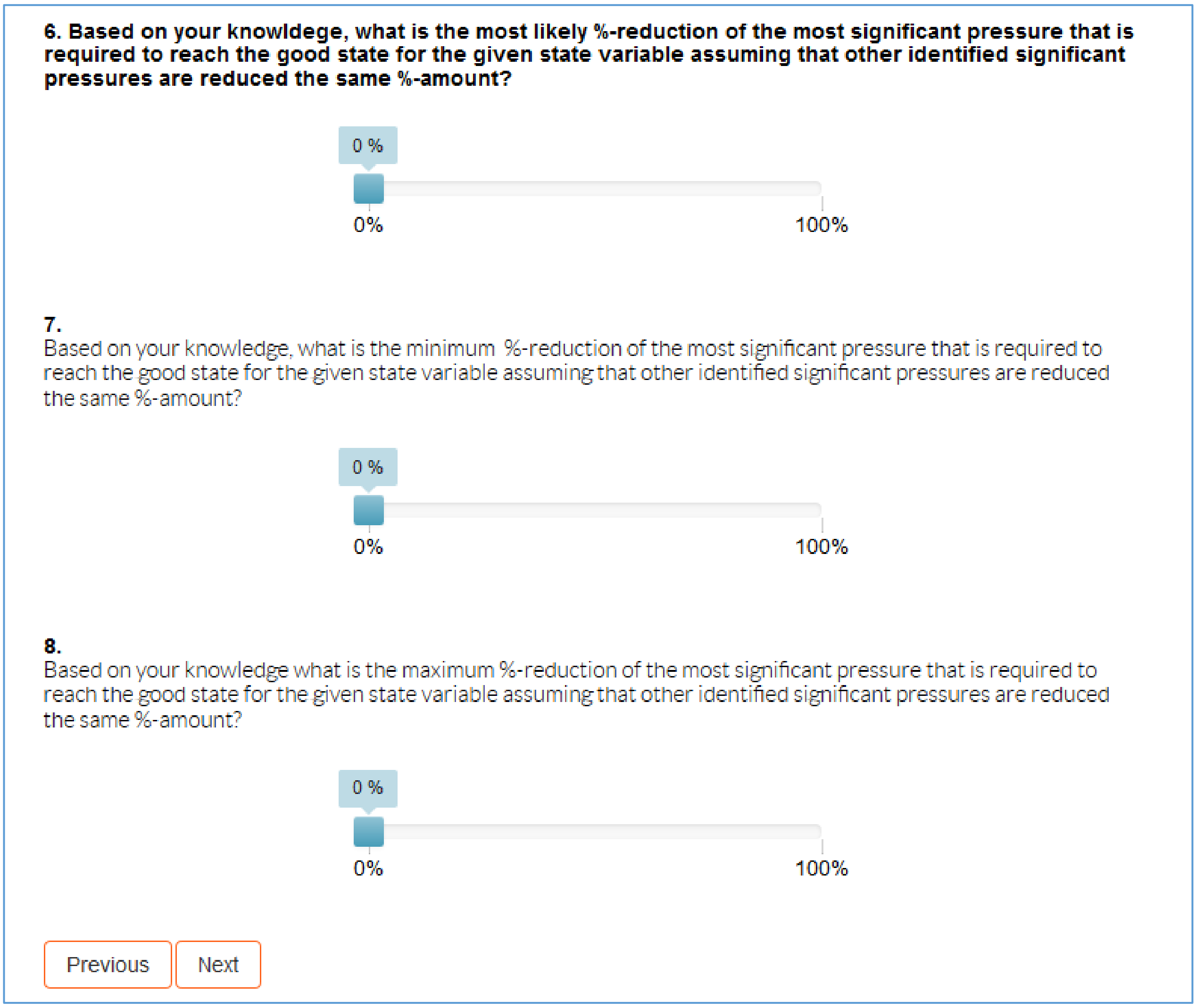


Figure 4. Question to define three-point estimates to assess the linkage between total pressure and state.

5. Comparison of BAU and GES, and sufficiency of measures (step 7)

The previous sections have outlined how to determine the expected pressure reduction distributions with existing measures, allowing for calculation of expected pressure reductions, confidence intervals and the probability to achieve specific percent (%) reduction in pressures. If we know a pressure target associated with good state and the current pressure level, we can estimate the pressure reduction required to reach a good state. Thus, for pressure variables that have a pressure target, we can assess whether the expected pressure reduction is sufficient to reach the target (i.e. if the expected pressure reduction from the existing measures is as large as the required pressure reduction), or estimate the probability of reaching a good state with the existing measures (=probability that given pressure reduction target is achieved from the distribution of the total pressure reduction with existing measures).

The cumulative distribution function of the total pressure reduction required to meet the good state $\boldsymbol{FS}_{k,s}$ is used to represent the probability of reaching a good state for different % reductions in total pressure affecting a state variable. If an expected value of total pressure reduction is applied to study how reductions in pressures increase the probability to reach good state, then cumulative distribution function $\boldsymbol{FS}_{k,s}$ can be used to define the expected probability to reach good environmental state. This also allows to estimate the required reductions in single significant pressures to reach good state, when other significant pressure reductions are known/unchanged. If the total pressure reduction is defined as a distribution, then the likelihood that the probability of reaching a good state is at least X% can be estimated.

When interpreting the results, one has to take into account the assumptions and generalizations that were made when defining the input distributions of activity-pressure, measure type effects and probability to reach good state, and the fact that these are based mainly on expert elicitations rather than empirical data.