The Research Council
of Norway

An evaluation of the Scientific Basis of the Traffic Light System for Norwegian Salmonid Aquaculture

Evaluation Committee, December 2021

The Research Council of Norway
Visiting address: Drammensveien 288
P.O.Box 564
NO-1327 Lysaker

Telephone: +47 22 03 70 00
post@rcn.no
www.rcn.no

Oslo, 21.12.2021

ISBN 978-82-12-03918-6 (pdf)

Evaluation Committee Members (in alphabetical order):

Kirstin Eliasen	Fiskaaling, Faroe Islands
David Jackson	Marine Biologist, Ireland
Anders Koed	DTU Aqua, Denmark
Crawford Revie	University of Strathclyde, United Kingdom (Chair)
Heather Anne Swanson	Aarhus University, Denmark
Jimmy Turnbull	University of Stirling, United Kingdom
Jarno Vanhatalo	University of Helsinki, Finland
André Visser	DTU Aqua, Denmark

Contents

Abbreviations used in this Report

- EvalComm – Evaluation Committee
- ExpGrp – Expert Group
- PA – Production Area
- SG – Steering Group
- TLS – Traffic Light System

Executive Summary

This document is the final report produced by an Evaluation Committee (EvalComm) set up under the auspices of the Research Council of Norway, at the request of the Ministry of Trade, Industry and Fisheries, to evaluate the scientific basis of the Traffic Light System (TLS) that is used to regulate the growth of the Norwegian salmon farming sector. The remit of the Committee was to focus primarily on two issues: (i) the **choice of scientific models and methods** (including their strengths and weaknesses, the handling of risk and uncertainty, and the quality of the assessments); and (ii) the extent to which the recommendations generated from the TLS **reflect the scientific evidence**.

The EvalComm was formally constituted in late 2020 and met over the course of a year to fulfil its remit, which included an Interim Note in July 2021 as well as this final report. In addition to meetings among EvalComm members (mostly using an on-line medium) there were also a number of interactions early in the process with members of the Expert and Steering Groups linked to the TLS, to clarify questions of operational process and to ensure that all relevant documentation was available to the EvalComm.

This documentation ran to over 1,000 pages, from nearly 100 documents, and as such it was not feasible, nor desirable, to comment on every aspect of the TLS, its various models, data sources and analyses. Instead the EvalComm chose to focus on a few key areas where it felt improvements could be made within the TLS and to create a series of recommendations to address weaknesses or limitations of the current approach. These range from the expert knowledge and data that are needed to create and parameterise the models within the TLS, to the interactions among those models, to the eventual interpretation of their outputs and transition into policy recommendations. Across all of these aspects of the TLS process, two cross-cutting themes stood out as being of critical importance in terms of ensuring the transparency and on-going legitimacy of the TLS approach: **uncertainty** and the role of **expert judgement** within the TLS.

It is clear that there are a number of places within the TLS where **uncertainty** is present. This may be the result of lack of access to adequate data or simply the inherent randomness associated with the biological and environmental variables being modelled. As such it is important to capture and properly quantify these sources of uncertainty. It is also critical to demonstrate the sensitivity of modelled outputs from the TLS to shifts around assumptions made about these uncertain parameters. There are also elements of uncertainty associated with the eventual risk assessments and consequent aquaculture policy advice being offered, all of which require that appropriate and easily understood language be used to communicate this uncertainty to a range of stakeholders.

It is also the case that the TLS depends on the inclusion of **expert judgement** at various points to address questions, ranging from the sources of data that should be included, to how various, and sometimes differing, sea lice risk indexes should be interpreted. There is sometimes a danger in systems as complex as the TLS, where the scientific evidence base is being given prominence, to downplay the nature or importance of such expert judgement within the process. The EvalComm believes this to be a mistake and a number of the recommendations relate to methods by which the process of expert elicitation can be captured in a more formal and transparent manner.

One of the challenges in a process as complex as the TLS is gaining a basic understanding of the key elements within the system and how these interplay with one another. To aid in this process a graphical overview of the TLS is provided early in the report. At various places in the report, the recommendations are ‘anchored’ to different locations within this graphic in an attempt to help the reader appreciate the relevance of a specific suggestion in the context of the overall TLS operation.

There are fifteen recommendations that relate to how various aspects of the TLS might be improved, a number of which have overlapping implications but arise in the context of different aspects of the EvalComm’s considerations. The final recommendation [R15] attempts to bring the other recommendations together by proposing a comprehensive framework for iterative review. The Driver-Pressure-State-Impact-Response (**DPSIR**) framework offers an approach for assessing the causes, consequences and responses to change in a complex adaptive system in a systematic way. It is the view of the EvalComm that DPSIR provides a framework that will support the continued relevance and value of the TLS, by ensuring it is constantly updated in line with new information and developments in best practice. The use of DPSIR should also aid in the communication of the processes underlying the TLS and its outcomes by a wider range of stakeholders.

1. Introduction

In this initial section of the report we provide some background as to the nature and purpose of the report, the process by which it was prepared, and the structure of the following sections.

Background and Remit

In late 2020 a group of international experts were set up at the request of the Norwegian Ministry of Trade, Industry and Fisheries, under the auspices of the Research Council of Norway, to evaluate the scientific basis of the Traffic Light System (TLS) that is used to regulate the growth of the Norwegian salmon farming sector. The composition of this Evaluation Committee (EvalComm), as we shall refer to it, is outlined below, along with the process through which this Final Report was produced.

Makeup of the Evaluation Committee (EvalComm) and key areas of expertise

The EvalComm consisted of eight scientists, drawn from seven research institutions in five different European countries (names and institutional affiliations can be found on the authorship list). The areas of expertise brought to the evaluation process by members of the Committee include: oceanographic modelling, fish biology, health and welfare, salmon farm management and regulation, host-parasite modelling, epidemiology, biostatistics, science communication, policy assessment, and science and technology studies.

Meeting Patterns and Development of Report

The EvalComm was formally constituted in late 2020 and between November 2020 and December 2021 conducted 14 on-line meetings (the minutes from which are available upon request). These meetings typically focused on specific aspects of the TLS, and towards the latter part of the process involved discussions on emerging sections of the draft final report. In June 2021, an Interim Report was prepared that consisted of, "…preliminary assessments and proposals for improvements".

In late November 2021, most of the EvalComm met in Copenhagen; those who could not attend in person participated via digital connection. During this meeting, the EvalComm drafted the final report based largely on the set of discussions that had taken place among members throughout the year. Different EvalComm members were responsible for drafting the various sections of this report. Subsequently, all members read the completed document and are in agreement on the set of recommendations that have been proposed.

Philosophy adopted during the review process

It was clear to the EvalComm from early in the process that comprehensive coverage of every part of the documentation, including all details of the data and modelling, would neither be possible or arguably helpful in conducting this review. As part of the process to ensure that relevant documents were given attention, a 'document matrix' (i.e. list of documents cross-tabulated with members of the EvalComm) was created; this resulted in a list of almost 100 documents and a total of over 1,000 pages of text. In light of this extensive documentation, not to mention the technical complexity of the many models that form component parts of the TLS, it was decided that focusing on key areas where the EvalComm concurred that potential improvements and/or extensions to the existing work of the Expert Group (ExpGrp) and Steering Group (SG) would be likely to yield the most significant reflections and possible improvements, would be a more productive approach than attempting to discuss all aspects of the scientific work.

This “key areas'' approach was outlined in an Interim Report that was completed in July 2021; including a note of our initial impressions as to the likely areas on which the EvalComm would focus. These areas coalesced during the initial six months of the EvalComm’s operation, which included a number of interactions with members of the SG and ExpGrp, and form the core of the analyses in Sections 3 and 4 of this report. Indeed, we noted in the Interim Report that, “we plan to focus on these areas during the next phase of our evaluation”; though as might be expected, a number of the recommendations have been amended and new suggestions added.

Before outlining these areas of focus, we would like to note that our assessment of the TLS is that it is probably the most sophisticated salmon risk assessment in operation around the globe in terms of the attempt to link research evidence to aquaculture policy. However, this does not mean that it would be the only practical or useful approach: other potentially simpler approaches could exist, as well. However, these alternatives are not considered in this document since the EvalComm did not interpret them to be part of its remit.

In addition, the EvalComm determined that evaluations and recommendations related to sea trout and Arctic char were outside the scope of the work it was to undertake. Salmon lice-induced mortality assessments have not been conducted for sea trout or Arctic char, meaning there is little data to currently evaluate. There has been a stated aim to expand the use of the TLS to cover both Arctic Char and Sea Trout. Finstad and colleagues proposed a method[1], where loss of marine habitat and loss of marine residence time could serve as sustainability indicators for sea trout and Arctic char in the traffic light system. However, the method has not yet been peer-reviewed, and sea trout and Arctic char are not yet included in the TLS. Thus, while the aspiration to extend the TLS to address potential impacts on other salmonid species was acknowledged to be important, it was not further evaluated by the EvalComm.

Structure of Report

Following on from these introductory comments on the process associated with carrying out the review and developing our report, we provide an overview of the TLS (Section 2). In addition to providing an outline of the constituent components that make up the TLS, this also allows us to introduce two of the most important, cross-cutting aspects of our review: the presence and importance of **uncertainty** within the TLS, as well as the role of **expert judgement** at various points within the framework. This is followed by one of the most significant parts of the report (Section 3) where we summarise what the EvalComm members view as the key areas where improvements could be made within the TLS. We then highlight a number of other important issues that may

---

[1] Finstad B, Sandvik AD, Ugedal O, Vollset KW, Karlsen Ø, Davidsen .G, Sægrov H (2019) **Appendix X** Sea trout (*Salmo trutta*) in the traffic light system – method proposal. **Sub-report from the ExpGrp**.

impact the future usefulness of the TLS (Section 4). The 'way forward' is then summarised (Section 5) by focusing on the importance of establishing various feedback loops to ensure that the TLS is fit for purpose and is delivering the conservation outcomes that it was designed to address. We propose the DSPIR framework for the development of an ongoing iterative performance analysis. Finally, the report ends (Section 6) with a set of overall conclusions and reflections on the scientific issues that are fundamental to the future potential of the TLS in Norway.

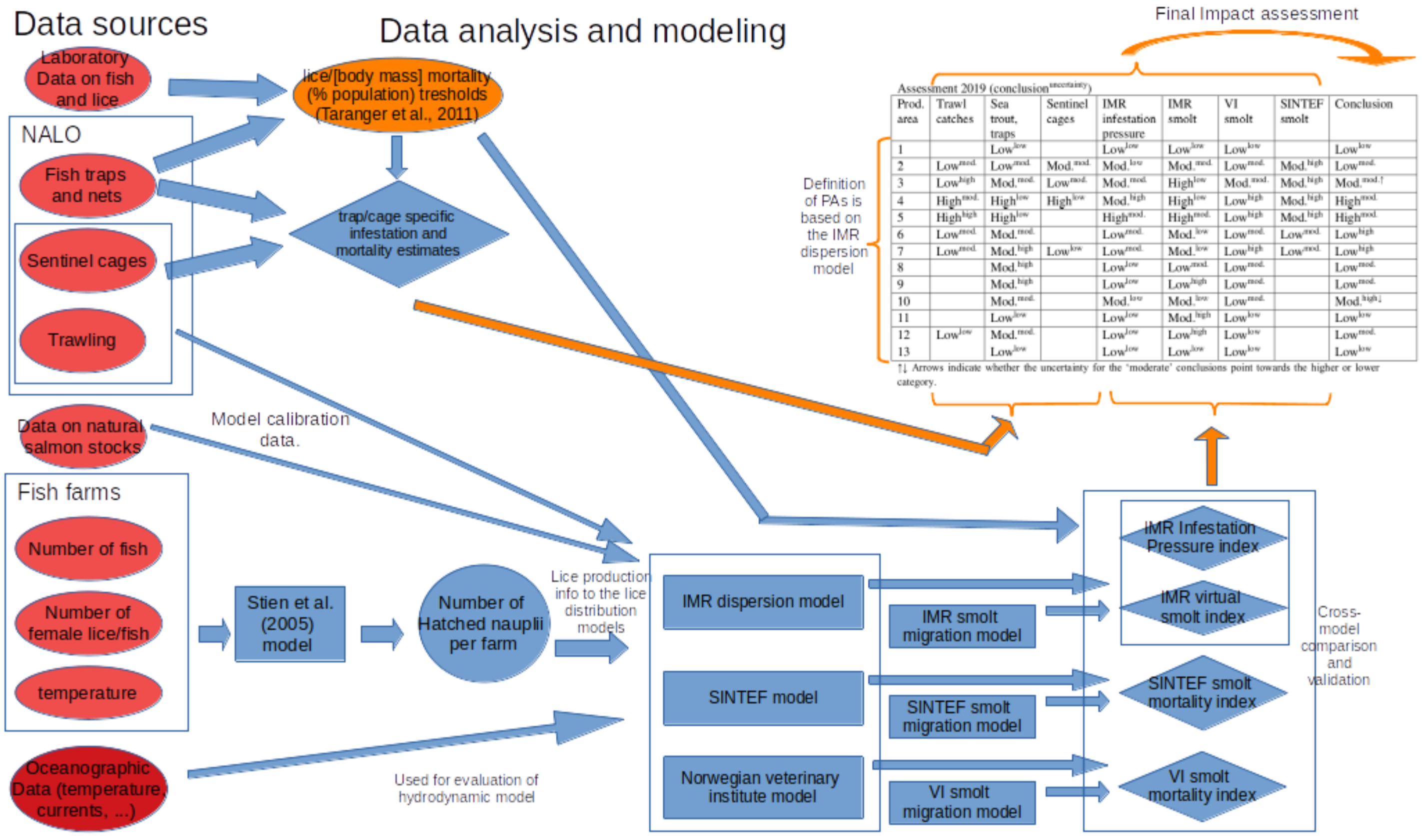
2. The Traffic Light System and points of focus in this report

The TLS assessment produced by the ExpGrp is a result of a complex process of integrating several models, data analyses and expert assessments. No simple overview that could be used to gain a holistic grasp of this process appeared to exist, so we produced a graphical representation (Figure 1) to aid our understanding. In creating Figure 1, we depicted the interdependencies between data sources and analytical steps as understood by the EvalComm, with the initial outline being modified in light of initial feedback from the ExpGrp. The figure is not in any sense intended to be a comprehensive description of all the details involved in the process; rather it summarises the key elements of the assessment process and highlights the interdependencies between different steps in the process.

In general, the overall risk assessment process in the TLS follows good scientific practice and tradition in these types of assessment. The impact assessment uses several sources of information, and uses process-based models to answer questions that are causal in nature (e.g. copepodid transmission). It proceeds with logical steps from empirical and theoretical knowledge towards the assessment end points (columns 2-8 in the assessment matrix; top-right corner in Figure 1). From the overall description of the assessment process, we can identify a few key elements.

First is the expert assessment related to sea lice mortality thresholds (essentially Taranger *et al.*, 2011). These thresholds are estimates for a percentage of the salmon population that die after being infected with a number of lice per body mass over the threshold value (orange oval at the top of Figure 1). The sea lice mortality thresholds are central for the whole TLS process since all sea lice impact indexes depend on them (see the blue arrows in Figure 1). The sea lice impact indexes also inform the individual assessment end points which are used to derive the final conclusions of TLS (arrow from columns 2-8 to column 9 in the assessment matrix). Hence, the seven complementary assessment end points are not mutually independent. There are also other sources of dependence among them, as depicted by arrows originating from a single source to multiple outputs in Figure 1, but the mortality threshold values are the single most important source of dependence between the assessment end points.

Second are the data (red ovals in Figure 1) that come in various forms and are used in the TLS assessment in multiple ways; these data are used to calibrate models, such as the sea lice dispersion and salmon migration models. Survey data from fish traps and nets, sentinel cages and trawling are also used to derive three sea lice impact indices (combined in the uppermost diamond in Figure 1) which then feed into three TLS assessment end points (columns 2-4 in the table in the upper right corner in Figure 1). Data are central to all subprocesses, analyses and conclusions in TLS for which



Assessment 2019 (conclusion^{uncertainty})

Prod. area	Trawl catches	Sea trout, traps	Sentinel cages	IMR infestation pressure	IMR smolt	VI smolt	SINTEF smolt	Conclusion
1		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}
2	Low$^{mod.}$	Low$^{mod.}$	Mod.$^{mod.}$	Mod.^{low}	Mod.$^{mod.}$	Low$^{mod.}$	Mod.^{high}	Low$^{mod.}$
3	Low^{high}	Mod.$^{mod.}$	Low$^{mod.}$	Mod.$^{mod.}$	High^{low}	Mod.$^{mod.}$	Mod.^{high}	Mod.$^{mod.\uparrow}$
4	High$^{mod.}$	High^{low}	High^{low}	Mod.^{high}	High^{low}	Low^{high}	Mod.^{high}	High$^{mod.}$
5	High^{high}	High^{low}		High$^{mod.}$	High$^{mod.}$	Low^{high}	Mod.^{high}	High$^{mod.}$
6	Low$^{mod.}$	Mod.$^{mod.}$		Low$^{mod.}$	Mod.^{low}	Low$^{mod.}$	Low$^{mod.}$	Low^{high}
7	Low$^{mod.}$	Mod.^{high}	Low^{low}	Low$^{mod.}$	Mod.^{low}	Low^{high}	Low$^{mod.}$	Low^{high}
8		Mod.^{high}		Low^{low}	Low$^{mod.}$	Low$^{mod.}$		Low$^{mod.}$
9		Mod.^{high}		Low^{low}	Low^{high}	Low$^{mod.}$		Low$^{mod.}$
10		Mod.$^{mod.}$		Mod.^{low}	Mod.^{low}	Low$^{mod.}$		Mod.$^{high\downarrow}$
11		Low^{low}		Low^{low}	Mod.^{high}	Low^{low}		Low^{low}
12	Low^{low}	Mod.$^{mod.}$		Low^{low}	Low^{high}	Low^{low}		Low$^{mod.}$
13		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}

↑↓ Arrows indicate whether the uncertainty for the 'moderate' conclusions point towards the higher or lower category.

Figure 1 A flow-chart of the impact assessment in the TLS for each PA. Red ovals represent data, model predictions (blue) or expert assessment (orange). Boxes represent models and diamonds represent sea lice impact indexes from model outputs or data analysis outputs. Blue arrows denote information flow from earlier analysis steps to later analysis steps. Orange arrows denote ExpGrp assessment processes that lead to complementary risk assessment end points (columns 2-8 in the assessment table) and to the final conclusions on each PA's status (column 9 in the assessment table). The assessment table at the top-right corner is copied from the ExpGrp 2019 report.

reason its quality (i.e. information content) and quantity (i.e. amount of information provided by data) significantly affect the final TLS conclusions.

Third are the sea lice dispersion models (blue rectangles in Figure 1) that simulate how sea lice produced in salmon farms spread in the water column and how they infect the migrating salmon. These are complex mechanistic, and deterministic, models that are calibrated and driven by vast amounts of observational and survey data. These models produce four sea lice impact indices (the diamonds in the lower right corner of Figure 1) which inform four TLS assessment end points (columns 5-8 in the table in Figure 1).

Fourth are the expert assessment steps (orange arrows in Figure 1) that are used in producing the actual risk assessment end points (columns 2-8 in the assessment matrix) from the individual sea lice impact indexes (blue diamonds) as well as the final conclusions (column 9 in the assessment matrix). This expert assessment step is the key concluding step in the whole TLS where the uncertainty in the complementary risk assessment indices is assessed and final conclusions made.

Based on the assessment of the overall structure of the TLS (Figure 1), there were six topics that the EvalComm considered important to explore. These topics, explored in Section 3, provide a framework to discuss key opportunities for improvement or development for the way in which the TLS is used in Norway. The topics are:

- **knowledge inclusion** (Section 3.1) What types of knowledge are included in the system building and evaluation process and who gets to determine these?
- **mortality threshold estimation** (Section 3.2) At what level of sea lice infestation do we consider a wild Atlantic salmon smolt to be at risk?
- **modelling framework** (Section 3.3) What implications flow from the choices made around the types of models developed and how these are integrated?
- **uncertainty estimation** (Section 3.4) How is the level of uncertainty at each step in the process evaluated?
- **uncertainty communication** (Section 3.5) How is uncertainty communicated to stakeholders and the end users of the assessments?
- **expert elicitation** (Section 3.6) How is expert assessment conducted in practice and how does this influence the final conclusions of the TLS assessment?

3. Key Areas for Development

3.1 Knowledge Inclusion

In its remit, the EvalComm was asked to assess "*the transparency and verifiability in the work of the Expert and Steering groups (documentation, publications etc.)*" alongside its technical work.

It is our impression that the ExpGrp and SG have shown a clear and admirable ambition to include a wide range of knowledge within the TLS. The instructions for the ExpGrp specifically suggest that it shall, "*be comprised of people from a broad range of backgrounds who possess expertise in the field and the ability to conduct an overall analysis of all available knowledge in order to arrive at a uniform assessment of salmon lice-induced wild fish mortality per production area*." The focus on a wide range of backgrounds and an openness to diverse knowledge sources is an important gesture that aligns with EU and UN statements about the importance of public participation in and access to science, as well as the importance of conducting "science in society."

In light of the ExpGrp and SG goals of (1) analysing all available knowledge, (2) having a committee comprised of people with expertise that position themselves to do so, and (3) ensuring transparent processes, we highlight several points for further strengthening these existing commitments.

Documentation of processes around knowledge inclusion

In response to our questions, the ExpGrp and SG indicated that they have conducted substantial outreach to allow members of various stakeholder groups and the general public to come forward with available knowledge relevant to salmon lice induced mortality. However, documentation and records of these activities, including processes of information solicitation and invitation to meetings, appears to be limited. We would like to recognize the efforts that the ExpGrp and SG appear to have made to undertake such activities. However, more substantial documentation of knowledge inclusion and the open, public solicitation of knowledge is important for transparency and legitimacy.

Such general documentation of knowledge inclusion practices should be coupled with an explicit policy on how submitted knowledge about salmon lice induced mortality is evaluated so that stakeholders and the public can better understand knowledge assessment processes.

Explicit statements on approaches to knowledge inclusion (who decides what is 'valid' and how?)

Due to limited information in the scientific peer-reviewed literature, it has sometimes been necessary for the ExpGrp to refer to a wider range of sources of information. This has included reports and other grey literature. In line with established practices for systematic reviews or meta-analyses, it is important that there is a policy for inclusion and exclusion of information in scientific assessments, in order to avoid bias. It is important that the ExpGrp adopts an approach that involves active reflection on decisions made around the inclusion or exclusion of information. This might usefully be linked to an overview of knowledge gaps in the existing data/information.

This also has relevance to the wider issue of knowledge inclusion. There is a great deal of experience and practical knowledge relating to farmed and wild salmon populations. The risk of using such information is that this may be anecdotal and biased. However, in the absence of a clear framework for inclusion or exclusion of sources of information, those with relevant knowledge, may well question why they were not consulted, or their opinion considered. It is clearly not the case that new knowledge is only generated by scientists; however, the processes by which knowledge generated by others (e.g. salmon farming industry, local communities, river management organisations, fishers) is incorporated into scientific modelling and other scientific assessment processes should be better documented and justified. Currently it is not possible to fully understand the framework that was used by the ExpGrp when deciding which sources to include.

We also suggest that attention is given to fostering a relatively equitable distribution of scientific capabilities and ability to conduct/benefit from scientific research across various stakeholder groups/communities. In particular, access to resources for formal research in relation to various topics that arise from ‘grey’ and/or local sources of knowledge, is important to consider as a part of the existing TLS commitment to assessing and engaging all available knowledge.

Improved communication of scientific results in forms accessible to a broad range of audiences

The value of providing some form of ‘systems overview’ (such as that shown in Figure 1) has already been noted, in the context of scientific critique and the identification of knowledge gaps. Elements such as this would be a useful addition to future reports to ensure that the processes leading up to the ultimate impact assessment and proposed actions can be understood by wider audiences. This issue of effective communication of scientific information to stakeholders and the general public, particularly with respect to uncertainty, is further considered in Section 3.4 of this report. It is also important that consideration is given to the nature of the TLS and how it is presented, so that audiences can more clearly understand how its *qualitative*, as well as its quantitative, elements form key parts of the overall process.

Recommendations

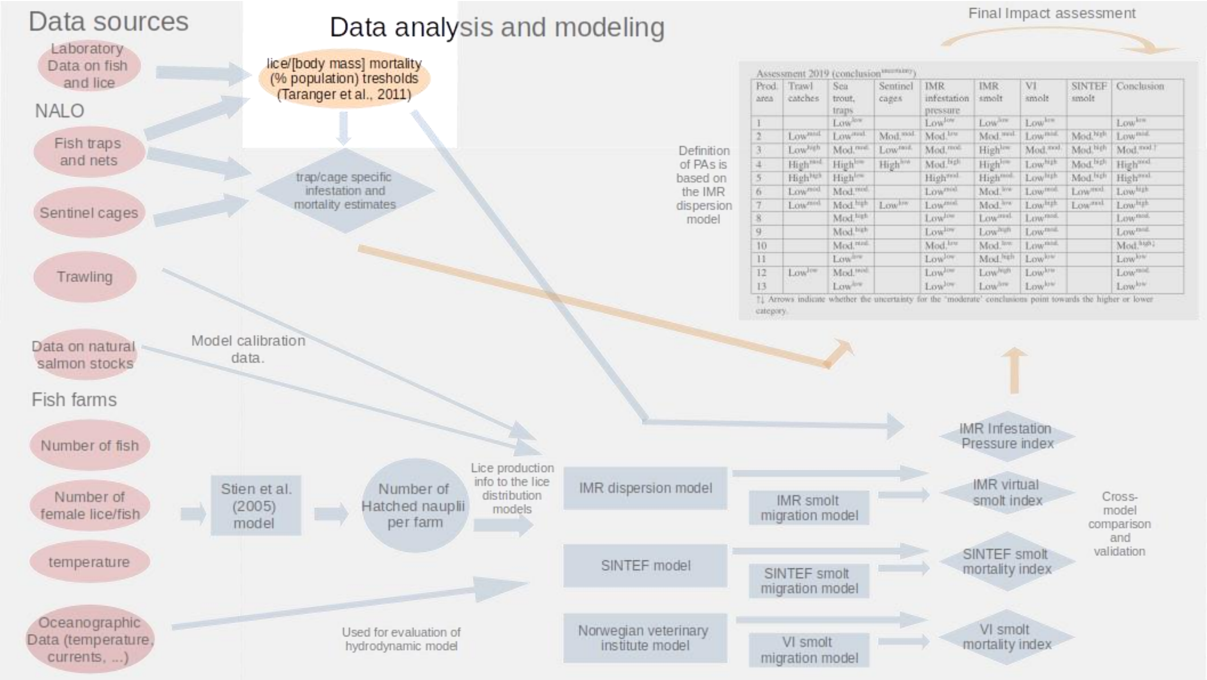
R1 We recommend more robust reporting of the **processes associated with knowledge inclusion** to ensure transparency and legitimacy.

R2 We recommend a **clear framework for inclusion or exclusion of sources of information** and an explicit policy on how submitted knowledge, especially reports and “grey” literature, will be evaluated.

R3 We recommend that further consideration be given to the **composition of the ExpGrp**, including the addition of expertise in areas such as scientific epistemology, knowledge inclusion, and science communication, to prompt continued reflection on such issues.

R4 We recommend that as part of the communication with stakeholders, proactive reflection be given to the **manner in which the TLS is presented** (i.e. is this a strictly quantitative system or more accurately a precautionary system with embedded quantitative approaches?).

3.2 Mortality Threshold Estimates



Assessment 2019 (conclusion^{uncertainty})

Prod. area	Trawl catches	Sea trout, traps	Sentinel cages	IMR infestation pressure	IMR smolt	VI smolt	SINTEF smolt	Conclusion
1		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}
2	Low$^{mod.}$	Low$^{mod.}$	Mod.$^{mod.}$	Mod.^{low}	Mod.$^{mod.}$	Low$^{mod.}$	Mod.^{high}	Low$^{mod.}$
3	Low^{high}	Mod.$^{mod.}$	Low$^{mod.}$	Mod.$^{mod.}$	High^{low}	Mod.$^{mod.}$	Mod.^{high}	Mod.$^{mod.\uparrow}$
4	High$^{mod.}$	High^{low}	High^{low}	Mod.^{high}	High^{low}	Low^{high}	Mod.^{high}	High$^{mod.}$
5	High^{high}	High^{low}		High$^{mod.}$	High$^{mod.}$	Low^{high}	Mod.^{high}	High$^{mod.}$
6	Low$^{mod.}$	Mod.$^{mod.}$		Low$^{mod.}$	Mod.^{low}	Low$^{mod.}$	Low$^{mod.}$	Low^{high}
7	Low$^{mod.}$	Mod.^{high}	Low^{low}	Low$^{mod.}$	Mod.^{low}	Low^{high}	Low$^{mod.}$	Low^{high}
8		Mod.^{high}		Low^{low}	Low$^{mod.}$	Low$^{mod.}$		Low$^{mod.}$
9		Mod.^{high}		Low^{low}	Low^{high}	Low$^{mod.}$		Low$^{mod.}$
10		Mod.$^{mod.}$		Mod.^{low}	Mod.^{low}	Low$^{mod.}$		Mod.$^{high\downarrow}$
11		Low^{low}		Low^{low}	Mod.^{high}	Low^{low}		Low^{low}
12	Low^{low}	Mod.$^{mod.}$		Low^{low}	Low^{high}	Low^{low}		Low$^{mod.}$
13		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}

↑↓ Arrows indicate whether the uncertainty for the 'moderate' conclusions point towards the higher or lower category.

Figure 1a Mortality threshold estimates are a central part of TLS feeding into all sea lice mortality indexes.

The mortality thresholds used in the current TLS are key factors in estimates of salmon lice induced mortality in wild salmonids (Figure 1a). The sources for the Mortality Threshold Estimates are the report Taranger *et al.* (2011) and a subsequent report, Taranger *et al.* (2012).

The authors of both Taranger *et al.* (2011) and the subsequent report, Taranger *et al.* (2012), made it clear that there was a lack of solid empirical evidence for the thresholds and that more research was necessary. As recently as 2019[2] the ExpGrp pointed to shortcomings in the methods used to determine the threshold values and recommended they be reviewed and evaluated again.

In 2019 a subgroup of the ExpGrp reviewed available data to determine whether there were grounds for changing the thresholds. They concluded that there was no basis for changing the limits proposed by Taranger *et al*. (2011 and 2012) and recommended further research including trials in nature.

In its Interim Note, the EvalComm raised the question of the need for a more complete scientific review of the basis for the Mortality Threshold estimates as outlined in Taranger *et al.* (2011). On the basis of the documentation provided to us, the latest document addressing this issue was a review

---

[2] Vollset KW, Nilsen F, Ellingsen I, Finstad B, Helgesen KO, Karlsen Ø, Sandvik AD, Sægrov H, Ugedal O, Qviller L, Dalvin S (2019) Assessment of salmon lice-induced wild fish mortality per production area in 2019. **Report from the ExpGrp**.

undertaken in 2019 by a subgroup[3] of the ExpGrp. The EvalComm is of the view that a solid empirical basis for the thresholds has not been provided to date and that such is required to underwrite key assessments arising from the TLS.

The TLS is interested in wild salmon stock levels and the vulnerability of smolts leaving rivers to sea lice infestation. The effects on salmon stocks will be a combination of intensity of lice infestation (level of challenge, duration of exposure, etc), with the impact of those lice. The impact is not just on mortality but overall reduction in fitness. It is difficult to envisage an empirical study that could provide any data allowing the effect of lice on wild smolts to be modelled more effectively. However, given that the current mortality thresholds have been in use since the inception of the TLS and in the absence of a solid empirical basis for the thresholds, we consider it essential that the mortality thresholds must be the focus of sensitivity analyses for overall TLS performance. The TLS is very sensitive to assumptions made in this part of the process, and yet empirical data appear to be largely unavailable. This represents a significant weakness in the TLS and its assessment end points.

There are substantial opportunities for improving the system by providing new documentation on the impact of salmon lice infestation on wild salmon. This documentation may focus on:

- Peer reviewed studies on mortality threshold and reduced fitness using a continuous approach (as opposed to a categorical) such as logistic regression, where the model variance could be integrated into the lice induced mortality estimates within the TLS.
- *In situ* studies on the effects of lice on salmon.
- Issues around over-dispersion and in general a heterogeneous fish/parasite population structure.

The absence of any documented reassessment of the validity of the mortality thresholds has the potential to undermine confidence in the operation of the system as a whole.

Recommendations

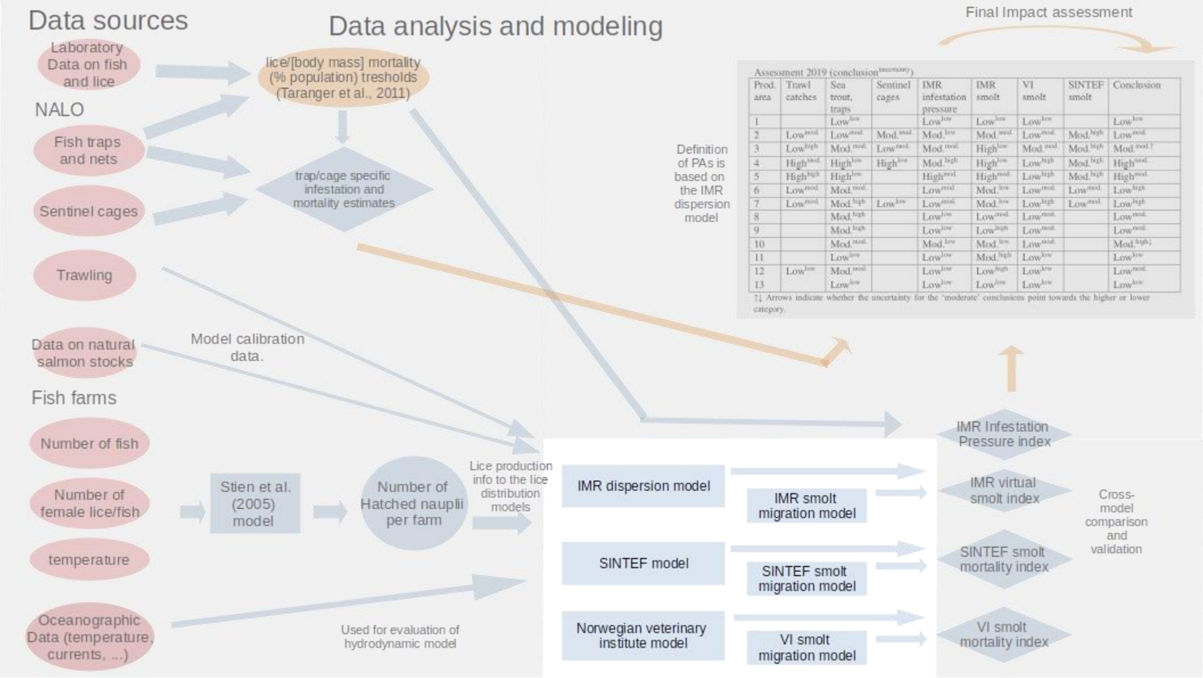
R5 We recommend that the **appropriateness of the mortality thresholds be reassessed regularly** in light of new information, with careful consideration around sensitivity analysis, as part of regular system performance reviews.

R6 We recommend that studies should be undertaken and peer reviewed, to **provide data on *in situ* effects of sea lice infestation** on wild salmon at an individual and population level.

---

[3] Karlsen Ø, Finstad B, Nilsen F (2019) An assessment of the mortality limits in use - updated with new information since 2012. **Appendix XI** in Assessment of salmon lice-induced wild fish mortality per production area in 2019. **Report from the ExpGrp**.

3.3 Modelling Framework



Assessment 2019 (conclusion^{uncertainty})

Prod. area	Trawl catches	Sea trout, traps	Sentinel cages	IMR infestation pressure	IMR smolt	VI smolt	SINTEF smolt	Conclusion
1		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}
2	Low$^{mod.}$	Low$^{mod.}$	Mod.$^{mod.}$	Mod.^{low}	Mod.$^{mod.}$	Low$^{mod.}$	Mod.^{high}	Low$^{mod.}$
3	Low^{high}	Mod.$^{mod.}$	Low$^{mod.}$	Mod.$^{mod.}$	High^{low}	Mod.$^{mod.}$	Mod.^{high}	Mod.$^{mod.\uparrow}$
4	High$^{mod.}$	High^{low}	High^{low}	Mod.^{high}	High^{low}	Low^{high}	Mod.^{high}	High$^{mod.}$
5	High^{high}	High^{low}		High$^{mod.}$	High$^{mod.}$	Low^{high}	Mod.^{high}	High$^{mod.}$
6	Low$^{mod.}$	Mod.$^{mod.}$		Low$^{mod.}$	Mod.^{low}	Low$^{mod.}$	Low$^{mod.}$	Low^{high}
7	Low$^{mod.}$	Mod.^{high}	Low^{low}	Low$^{mod.}$	Mod.^{low}	Low^{high}	Low$^{mod.}$	Low^{high}
8		Mod.^{high}		Low^{low}	Low$^{mod.}$	Low$^{mod.}$		Low$^{mod.}$
9		Mod.^{high}		Low^{low}	Low^{high}	Low$^{mod.}$		Low$^{mod.}$
10		Mod.$^{mod.}$		Mod.^{low}	Mod.^{low}	Low$^{mod.}$		Mod.$^{high\downarrow}$
11		Low^{low}		Low^{low}	Mod.^{high}	Low^{low}		Low^{low}
12	Low^{low}	Mod.$^{mod.}$		Low^{low}	Low^{high}	Low^{low}		Low$^{mod.}$
13		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}

↑↓ Arrows indicate whether the uncertainty for the 'moderate' conclusions point towards the higher or lower category.

Figure 1b The modelling framework comprised of IMR, SINTEF and NVI lice dispersion models and smolt migration models. These are used to calculate four separate sea lice mortality and infestation pressure indexes.

Several of the sea lice impact indexes that go into the TLS come by way of three modelling systems: the IMR Model System, the NVI Risk Model, and SINTEF's SINMOD model.

The **IMR (Institute for Marine Research) model** is composed of three components. Firstly, it includes a numerical hydrodynamic model forced by wind, freshwater discharge, and tides that is informed through near real-time data assimilation. This produces high resolution simulations of circulation and relevant water properties along the Norwegian coast.

This feeds a Lagrangian dispersion model coupled to a life history model of drifting sea lice. The life history component simulates the growth, mortality and behaviour of sea lice nauplii and copepodids as a function of the temperature and salinity they encounter. Input data for this sea lice model comes from reported fish farm infection rates and extrapolated through an empirical formulation (Ådlandsvik 2017). The most sensitive aspect of this modelling component is in the vertical behaviour of nauplii and copepodids. Given the high vertical variation of horizontal currents in fjord systems, small variations in vertical positioning can have large implications on subsequent drift and dispersion patterns. The final component is a model for the migration of salmon smolt from rivers, through the farming region to the open sea.

The aggregated outputs (sea lice impact indexes; blue diamonds in Figure 1b) produced by this modelling system that feed into the decision-making process (orange arrows in Figure 1b) are:

- *Infestation pressure* (ROC: relative operating characteristic): These estimates are produced from maps of infestation pressure (the probability that a resident salmon will become infected). These maps are calibrated against sentinel cage observations (see Section 4). In essence, these maps represent an informed fit of sentinel cage observations along the entire Norwegian coastal system. To represent each PA, the weighted infestation pressure in terms of areal extent is calculated to produce an index (blue diamond "IMR infestation pressure index" in Figure 1). There is some degree of expert judgement here as the total area of each PA is somewhat subjective.

- *Virtual post-smolt infestation*: A random walk is used to simulate the passage of smolt from rivers to the open sea[4]. Numbers of lice per smolt are estimated from infestation pressure maps depending on timing in the migration season[5], and transit time through different regions. The infestation per fish is calibrated against trawl data which genetically resolves the river of origin for a limited number of catchments. Mortality rate is taken as a simple function of lice per smolt based on mortality thresholds (see 3.2 Mortality Threshold Estimates).

The model system is validated against infection counts in sentinel cages, fish traps and trawls (Section 4). As more years of data accumulate, this validation and subsequent calibration is being updated. The model system has gone through at least one major re-write to facilitate an improved particle tracking algorithm, which has required a re-calibration. In addition, the sensitivity to vertical positioning is in a process of continual updates, as new experiments and observations become available.

The **NVI (Norwegian Veterinary Institute) Risk Model** is based on much the same reasoning as the IMR model above; fish farm infection counts are used to estimate egg production rates and a life history model is used to estimate growth rates and mortality as eggs develop to nauplii and copepodids. Dispersion, however, is not estimated from a hydrodynamic model, but rather as a diffusive process estimated from long term observations. These produce estimated infestation pressure maps that are calibrated against sentinel cage observations (see Section 4), with linear regression models providing confidence limits. As the number of years of observations increases, the calibration of this pivotal step is being updated.

The time that seaward migrating smolts spend in different regions, as they emerge from various water courses, provides an estimate of infestation rates. These infestation rates are calculated three times during the migration period and are converted to mortality based on mortality thresholds (3.2

---

[4] Johnsen IA (2020) Seaward migration of virtual postsmolts. **Appendix VI** in Assessment of salmon lice-induced wild fish mortality per production area in 2020. **Report from the ExpGrp**.

[5] Vollset KW, Lennox R, Ugedal O, Sægrov H (2020) New model for outmigration of salmon smolt. **Appendix IX** in Assessment of salmon lice-induced wild fish mortality per production area in 2020. **Report from the ExpGrp**.

Mortality Threshold Estimates). Water course-resolved mortality rates weighted by potential smolt production are aggregated per PA to provide the risk assessment.

The **SINTEF model (SINMOD)** is a modelling system composed of hydrodynamic, life history and smolt sub-modules. The hydrodynamic model is nested, meaning it takes a low-resolution regional model covering the entire North Atlantic to drive a succession of smaller scale, higher resolution models eventually down to the scale of the Norwegian coastal system. It is subject to similar forcing as the IMR model, has a higher resolution but is used in a limited number of PAs (PA2 to PA7). Input to the model is similar to the other models; reported infestation rates in farms are extrapolated to a daily basis to provide an estimate of time and locations of egg sources. A life history model following Stien *et al*. (2005) provides a structured population model of the sea lice dispersive stages. Unlike the IMR model, SINMOD is a Eulerian model, meaning that it estimates advection and diffusion of sea lice between grid boxes. It also includes vertical behaviour of sea lice as they position themselves in the water column in response to temperature, light and salinity. The infestation pressure on migrating smolts is estimated using a particle tracking model simulating the drift and active swimming of smolts. Infestation per smolt is estimated using an encounter rate model (volume swept per smolt times sea louse concentration times infestation probability). Calibration of the infestation rate in SINMOD is in terms of the observed and simulated frequency distribution of attached lice, to give the frequency distribution of mortality.

Each of these model systems are run independently to provide four different sea lice impact indexes that go into the expert assessment process. Individually, and taken as a whole, the modelling systems represent a state-of-the-art network approach to simulate the impact of sea lice in the area of coastal salmon farms. The different methodologies of the modelling systems provide the ExpGrp with independent metrics that can be counted as an important, added value dimension to the decision-making process.

It is a strength that these modelling systems are in a constant state of re-appraisal, and that as new information becomes available, it is incorporated into the numerical models. A case in point is the re-evaluation of vertical positioning of sea lice, and what this means for predicted dispersion in the IMR model. An important feature of modelling systems, demonstrated here, is that new numerical descriptions should be rigorously tested through hindcasts. This builds not only confidence in the models themselves, but also ensures a continuity of results that can be used to evaluate system performance.

Although their methodologies are different, all three systems rely on much of the same input data (farm infestation rates, sentinel cages, trawls etc.; see blue arrows from data sources (red ovals) in Figure 1) and many of the same descriptions of, for instance sea lice life history (e.g. Stien *et al*. 2005) and lice induced mortality in salmon (Taranger *et al*. 2011; orange oval in Figure 1). It is important that these descriptions do not become entrenched in the traffic light assessment process. More importantly, while the models themselves are state-of-the-art, model products are heavily reliant on calibration data, some of which demonstrate high levels of variability and uncertainty (e.g. sentinel cages). This reliance on calibration data potentially impairs the quality of model products. Further, the perception that models are advanced computational methods, can lead to a perceived over-confidence in their results. Model results should be presented as much as possible with confidence intervals through formal uncertainty quantification.

The three modelling systems each provide a wealth of information that is aggregated into a few metrics for assessment purposes. It is clear that each is embedded in a research environment that utilizes this information for scientific purposes. Within this context, there is also the possibility that these models can run virtual experiments. This is already done in a simple sense, through sensitivity studies on parameter values. Extending this concept, there is also the opportunity to run scenarios with variable numbers of salmon farms. While we do not advocate that this become an element of the TLS, it is an area of model development that can add to their scientific impact.

One of the major threats to the modelling systems is that they can potentially fail to capture long term trends both local (warming, freshening) and external (changing migration patterns in wild stocks). Perhaps more importantly, they can fail to capture rapid transitions (e.g. regime shifts, tipping points) and their drivers (e.g. marine heat waves). It is the nature of successful model architecture that they tend to avoid instabilities. Further, these models become unreliable as they are pushed outside of their calibration envelopes. While not of immediate concern, rapidly changing climate as witnessed particularly in Arctic areas, could precipitate a system change that these models would fail to capture.

In summary, within the constraints of the quality and availability of observations and current knowledge, the system of models in the TLS are state-of-the-art. As referred to above, the models use data which has vulnerabilities and, in some cases, it is not obvious how such vulnerabilities might be resolved.

The product of these models may appear highly precise and quantitative, but their output has to be used with full cognizance of their limitations, and care must be taken when explaining the model outputs to non-expert audiences (see 3.1 Knowledge inclusion and 3.6 Expert elicitation).

Recommendation:

R7 We recommend that rather than investing effort in refining the models themselves, the models should be used to map out **sensitivity** and **identify sources of uncertainty** that can be most easily addressed with additional observations.

3.4 Uncertainty Estimation

The final assessment produced by the ExpGrp is a result of a rather complex process of integrating several modelling and data analysis steps (Figure 1). Hence, for the sake of transparency and clarity, the full picture of the assessment process should be clearly described. It is our observation that the ExpGrp reports lack a clear description of uncertainty estimation and as a result it is difficult to understand what data feeds into what model, or how different model results and predictions are connected to one another. Specifically, risk assessment reports should include sensitivity analyses of the **individual model components** as well as on the **process of integrating the results of these sub-models** into the final assessment. The sensitivity analyses of individual model components in the TLS are reported to varying degrees either in the ExpGrp reports or scientific publications referred to in the ExpGrp reports.

We have examined in some detail the process by which uncertainty is propagated through the different sub-models to the final impact assessment. Based on the ExpGrp reports and interviews with the ExpGrp members, the final impact assessment is carried out in an ExpGrp meeting based on: (i) the results from intermediate impact assessments, i.e. sea lice impact indexes (blue diamonds in Figure 1), and (ii) questionnaires that summarize the uncertainty in and the reliability of these sea lice indexes, for each of the PAs[6]. The sea lice impact indexes include the smolt mortality indexes of IMR, VI, and SINTEF, as well as the sea lice impact indexes calculated from the sentinel cages, fish traps and netting data (see Figure 1). This process which leads to the final impact assessment is neither fully transparent nor rigorously reported and as such leaves room for criticism and doubt concerning the relative contributions of different sub-models and data to the final uncertainty assessment (see also Section 3.6 on Expert Elicitation).

While we acknowledge that due to the high complexity of the impact assessment process, formal uncertainty quantification is difficult in practice, it is possible in principle, and there exist practical tools to tackle aspects of such tasks. First, the definition of uncertainty and its quantification (preferably, using probability) should be harmonized across the different sub-models. Second, the process represented in Figure 1 is Markovian in nature (each step in the process is dependent only on the immediately preceding steps), as such it should be possible to carry out formal qualitative, and likely also quantitative, uncertainty propagation in a step-wise manner.

Recommendation:

R8 We recommend that model, data analysis and expert assessment results be presented with confidence intervals as much as possible to avoid a misplaced perception of their accuracy. **Proper uncertainty quantification for the models** would increase the reliability of the conclusions based on them.

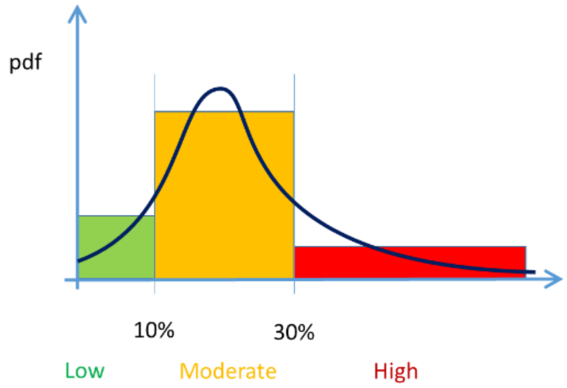
---

[6] Questions used when assessing the production areas. **Appendix XII** in Assessment of salmon lice-induced wild fish mortality per production area in 2020. **Report from the ExpGrp**.

3.5 Uncertainty Communication

The approaches currently adopted within the TLS for **reporting uncertainty** do not reflect either best scientific practice or plain-language communication understandable to policy makers and other stakeholders. This is not an issue that is unique to the TLS, and has been revisited many times in recent years in association with issues ranging from climate change to public health.

The TLS is predicated on categorizing PAs according to the expectation of salmon-lice induced mortality into one of three tiers: Low, Moderate or High. While there is much to be said about sources of uncertainty, for any prediction, there will be a probability of a PA being in one of these 3 states. At issue is how this distribution of probability can be best communicated. Simply seen, it is conveying the information in the probability density function (see example opposite), where the shape of the function itself is determined by all kinds of input, from models, observations and expert assessment. The output of the procedure is to assign a category (Low, Moderate, High) to a PA and relate an uncertainty of this assignment.



The current practice[7] focuses on the uncertainty in assigning the correct traffic light category to a given PA. The criteria[8] used are:

- High uncertainty = the probability that the category is correctly defined exceeds 50%, but there is a 35–49.9% probability that it is either lower or higher.
- Moderate uncertainty = the probability that the category is correctly defined exceeds 50%, but there is a 20–34.9% probability that it is either lower or higher.
- Low uncertainty = the probability that the category is correctly defined exceeds 50%, but there is a 0–19.9% probability that it is either lower or higher.

This uses the concept of the preponderance of probability, and is applied assuming that the three category assignments can, for all intents and purposes, be whittled down to two. There are some issues concerning this particular protocol, in particular its scientific rigor and its plain-language interpretation.

Firstly, given that the probability of the PA being in one of the three categories is 100%, this seems to simply boil down to High, Moderate and Low uncertainty being related to the probability of the PA

---

[7] Vollset KW, Nilsen F, Ellingsen I, Finstad B, Helgesen KO, Karlsen Ø, Sandvik AD, Sægrov H, Ugedal O, Qviller L, Dalvin S (2019) Assessment of salmon lice- induced wild fish mortality per production area in 2019. **Report from an ExpGrp**. P10.

[8] Memo on the description of uncertainty in the main conclusions for each production area (15th Nov 2019)

being in a specific category lying between 80%-100%, 65%-80% and 50%-65% respectively. Assigning the category can be done hierarchically according to rules of the form:

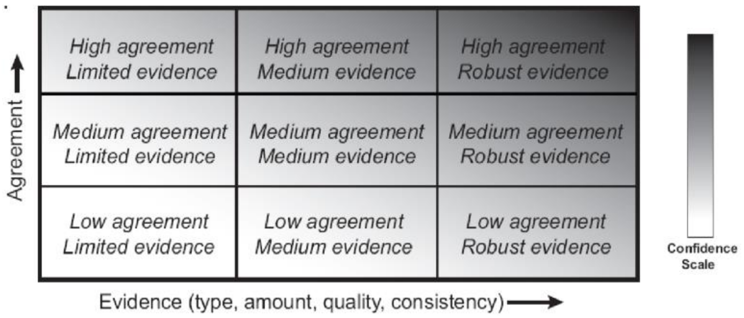
- P3 (Probability of >30% mortality) exceeds 50% then category High with uncertainty according to High uncertainty (50%<P3<66%), Moderate uncertainty (66%<P3<80%), Low uncertainty (80%<P3<100%)
- P1 (Probability of <10% mortality) exceeds 50% then category Low with uncertainty according to High uncertainty (50%<P1<66%), Moderate uncertainty (66%<P1<80%), Low uncertainty (80%<P1<100%)
- P2 (100% – P3 – P1) exceeds 50% then category Moderate with uncertainty according to High uncertainty (50%<P2<66%), Moderate uncertainty (66%<P2<80%), Low uncertainty (80%<P2<100%)

Given there are three categories, there is also a fourth uncertainty where neither P1, P2 nor P3 exceed 50%. On a 3-tier scale this is technically indeterminate, although a practical classification would be Moderate with Very High uncertainty.

These technical aspects aside, there is now the question as to how to communicate this uncertainty to policy makers and the public. This issue has been taken up in several advisory bodies working at the science-policy interface.

The IPCC (Intergovernmental Panel on Climate Change) has developed a set of protocols that have tried to standardise language regarding uncertainty. They divide this into two parts and convey the concepts of both “confidence” and “uncertainty” (or “likelihood”). Regarding this, they rely on two metrics for communicating the degree of certainty in key findings[9]:

- **Confidence** in the validity of a finding, based on the type, amount, quality, and consistency of evidence (e.g., mechanistic understanding, theory, data, models, expert judgment) and the degree of agreement. Confidence is expressed qualitatively.
- Quantified **measures of uncertainty** in a finding expressed probabilistically (based on statistical analysis of observations or model results, or expert judgment).



High agreement *Limited evidence*	*High agreement* *Medium evidence*	*High agreement* *Robust evidence*
Medium agreement *Limited evidence*	*Medium agreement* *Medium evidence*	*Medium agreement* *Robust evidence*
Low agreement *Limited evidence*	*Low agreement* *Medium evidence*	*Low agreement* *Robust evidence*

Confidence relates to both the quality of evidence and the consistency of agreement; high agreement and robust evidence implies high confidence, whereas poor evidence and low agreement implies low confidence.

---

[9] Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties, IPCC Cross-Working Group Meeting on Consistent Treatment of Uncertainties, Jasper Ridge, CA, USA, 6-7 July 2010

In terms of expressing uncertainty, the IPCC adopted relatively simple language to convey how likely a prediction or observation is; likely, very likely, virtually certain, etc. These are assigned specific statistical probability intervals to provide scientific rigor (see opposite).

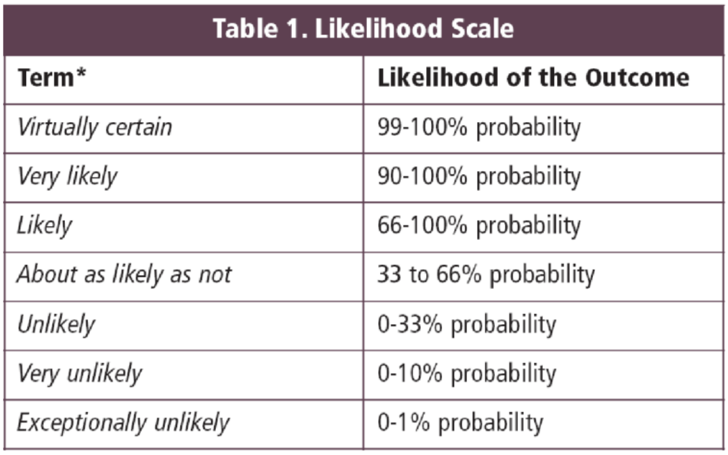


Table 1. Likelihood Scale

Term*	**Likelihood of the Outcome**
Virtually certain	99-100% probability
Very likely	90-100% probability
Likely	66-100% probability
About as likely as not	33 to 66% probability
Unlikely	0-33% probability
Very unlikely	0-10% probability
Exceptionally unlikely	0-1% probability

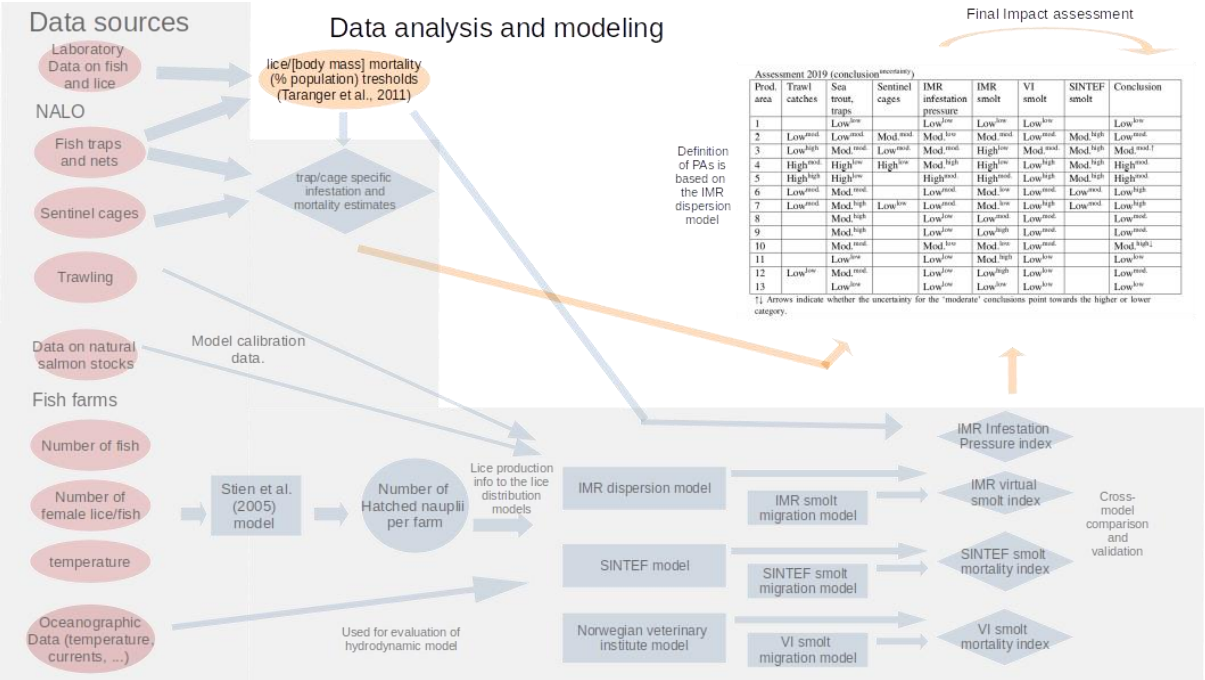
It can be debated as to exactly how well and in what circumstance these categories mesh with scientific and public perceptions, but this has become a benchmark that is gradually gaining wider acceptance and uptake. It would be worthwhile investigating how this might be used with the TLS to describe the statements around the likelihood of a particular PA being in the Low, Moderate or High 'traffic light' category.

Recommendations

R9 We recommend that a **more transparent and rigorous reporting process** around system sensitivity and uncertainty be provided.

R10 We recommend that the use of **more easily understood language** be explored when conveying the confidence and uncertainty associated with TLS assessments. A particular concern here is how these aspects of the TLS are communicated beyond a scientific audience to policy makers, other stakeholders and the general public. The ExpGrp should look to the IPCC as an example of a relatively successful protocol.

3.6 Expert Elicitation



Assessment 2019 (conclusion^{uncertainty})

Prod. area	Trawl catches	Sea trout, traps	Sentinel cages	IMR infestation pressure	IMR smolt	VI smolt	SINTEF smolt	Conclusion
1		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}
2	Low$^{mod.}$	Low$^{mod.}$	Mod.$^{mod.}$	Mod.^{low}	Mod.$^{mod.}$	Low$^{mod.}$	Mod.^{high}	Low$^{mod.}$
3	Low^{high}	Mod.$^{mod.}$	Low$^{mod.}$	Mod.$^{mod.}$	High^{low}	Mod.$^{mod.}$	Mod.^{high}	Mod.$^{mod.\uparrow}$
4	High$^{mod.}$	High^{low}	High^{low}	Mod.^{high}	High^{low}	Low^{high}	Mod.^{high}	High$^{mod.}$
5	High^{high}	High^{low}		High$^{mod.}$	High$^{mod.}$	Low^{high}	Mod.^{high}	High$^{mod.}$
6	Low$^{mod.}$	Mod.$^{mod.}$		Low$^{mod.}$	Mod.^{low}	Low$^{mod.}$	Low$^{mod.}$	Low^{high}
7	Low$^{mod.}$	Mod.^{high}	Low^{low}	Low$^{mod.}$	Mod.^{low}	Low^{high}	Low$^{mod.}$	Low^{high}
8		Mod.^{high}		Low^{low}	Low$^{mod.}$	Low$^{mod.}$		Low$^{mod.}$
9		Mod.^{high}		Low^{low}	Low^{high}	Low$^{mod.}$		Low$^{mod.}$
10		Mod.$^{mod.}$		Mod.^{low}	Mod.^{low}	Low$^{mod.}$		Mod.$^{high\downarrow}$
11		Low^{low}		Low^{low}	Mod.^{high}	Low^{low}		Low^{low}
12	Low^{low}	Mod.$^{mod.}$		Low^{low}	Low^{high}	Low^{low}		Low$^{mod.}$
13		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}

↑↓ Arrows indicate whether the uncertainty for the 'moderate' conclusions point towards the higher or lower category.

Figure 1c Expert judgement is an integral part of TLS in several places. However, it is most apparent in the formulation of mortality thresholds and in the final assessment process by ExpGrp where the risk assessment table and final conclusions are done (top-right corner).

Large parts of the sea lice impact assessment in the TLS relies on expert judgement. This is most clearly visible in the final steps of the assessment where the ExpGrp's overall conclusions are made, and the conclusions table compiled (orange arrows in Figure 1c). Even though the final conclusions are based on data (red ovals) and extensive modelling and data analysis steps (blue arrows and boxes), all these earlier steps are assessed by experts to arrive at the final conclusions. Expert judgement is also strongly present in setting up mortality limits, evaluation of individual assessment outcomes (columns 2-8 of the final conclusions table) and in judgements around modelling and statistical analyses. Expert judgement is a natural part of any risk assessment and as such is a natural ingredient of the TLS.

The ExpGrp report is explicit in its use of expert judgement in the final conclusions step. The practice of having humans making the final conclusions enhances the credibility and ethical validity of systems such as the TLS. Conclusions should not be blindly based on model outputs but require an overall assessment by a group of experts in the field, as is the current practice in the TLS.

The process and methods to compile expert judgement in practice is often referred to as *expert elicitation*. Apart from the modelling and statistical analysis steps, which are reported using standard scientific practice, the practical implementation of expert elicitation, and the process used to derive conclusions from it, is poorly documented in the ExpGrp reports and their appendixes. The roles of the ExpGrp members in the process of making the final conclusions are not clearly described. The process and logic of eliciting uncertainties in the final conclusions is especially problematic (Sections 3.4 and 3.5). Moreover, it seems that the expert elicitation process leading to the final conclusions does not follow any accepted best practice standard, such as those laid out in risk assessment fields (e.g. European Food Safety Authority, 2014), statistics (e.g. O'Hagan, 2019) and climate change research (e.g. IPCC, 2010). Hence, the chosen method would need to be described and justified before its validity could be accepted.

The ExpGrp contains expertise from the key fields related to sea lice and salmon farming, which gives a good opportunity to make excellent final conclusions. The model outputs and infestation rates derived from the sentinel cages, trawl catch, and sea trout traps provide a quantitative knowledge base on which to base the final conclusions of the ExpGrp. This gives a good starting point to make structured, uncertainty-explicit expert elicitation. Research on expert elicitation methods is active, and many fields of risk assessment (such as food safety and climate change) have developed and tested carefully designed elicitation protocols, which could be adapted and further developed for the purposes of the TLS.

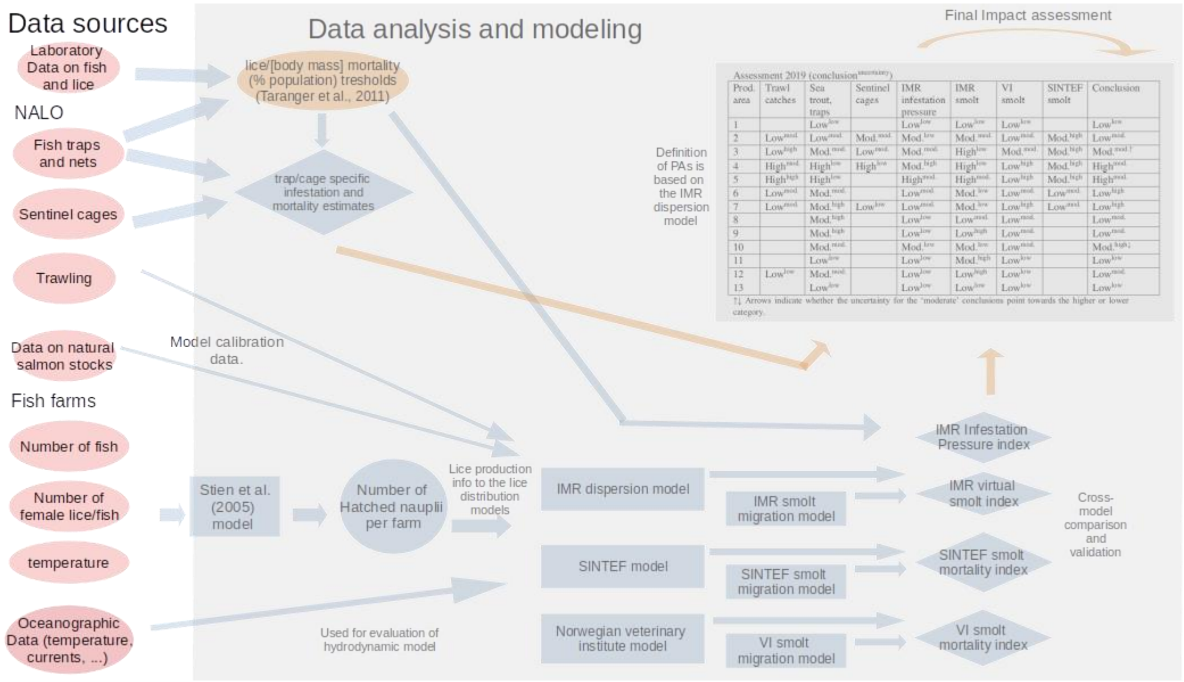
Eliciting expert knowledge (including overall conclusions from risk modelling such as the TLS) carefully and scientifically is not a simple task. Psychologists have identified numerous ways in which naive questioning can promote cognitive biases in the experts' judgements. To elicit expert knowledge as objectively as possible, the elicitation process needs to be structured, and preferably facilitated by experienced elicitor(s), so as to minimize such biases. This is especially important in situations where experts are asked to assess probabilities (i.e. uncertainties) and when they come from different backgrounds, such that common language and terminology might be missing. Because expert elicitation in the TLS is poorly documented and does not follow standards in the field, there is a significant threat that the assessment of mortality limits and final conclusions of the ExpGrp are prone to cognitive biases.

Recommendation:

R11 We recommend that the expert assessment processes within the TLS should be **described and justified in light of accepted best practice standards** that exist in the area of expert elicitation.

4. Other Considerations / Data Vulnerabilities

There are vulnerabilities in some of the data sources that underpin much of the TLS (depicted as red circles in Figure 1d). While there is the potential for some of these data sources to be refined or improved, the contribution of such improvements to the overall TLS performance would need to be considered before recommending substantial changes. In some cases, the data are inherently unreliable; in others, while the data might be refined, it is unlikely these refinements will improve the overall performance of the TLS, given the expert judgement process that leads to the eventual classification of red, yellow or green.



Assessment 2019 (conclusion^{uncertainty})

Prod. area	Trawl catches	Sea trout, traps	Sentinel cages	IMR infestation pressure	IMR smolt	VI smolt	SINTEF smolt	Conclusion
1		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}
2	Low^{mod}	Low^{mod}	Mod.^{mod}	Mod.^{low}	Mod.^{mod}	Low^{mod}	Mod.^{high}	Low^{mod}
3	Low^{high}	Mod.^{mod}	Low^{mod}	Mod.^{mod}	High^{low}	Mod.^{mod}	Mod.^{high}	Mod.$^{mod\uparrow}$
4	High^{mod}	High^{low}	High^{low}	Mod.^{high}	High^{low}	Low^{high}	Mod.^{high}	High^{mod}
5	High^{high}	High^{low}		High^{mod}	High^{mod}	Low^{high}	Mod.^{high}	High^{mod}
6	Low^{mod}	Mod.^{mod}		Low^{mod}	Mod.^{low}	Low^{mod}	Low^{mod}	Low^{high}
7	Low^{mod}	Mod.^{high}	Low^{low}	Low^{mod}	Mod.^{low}	Low^{high}	Low^{mod}	Low^{high}
8		Mod.^{high}		Low^{low}	Low^{mod}	Low^{mod}		Low^{mod}
9		Mod.^{high}		Low^{low}	Low^{high}	Low^{mod}		Low^{mod}
10		Mod.^{mod}		Mod.^{low}	Mod.^{low}	Low^{mod}		Mod.$^{high\downarrow}$
11		Low^{low}		Low^{low}	Mod.^{high}	Low^{low}		Low^{low}
12	Low^{low}	Mod.^{mod}		Low^{low}	Low^{high}	Low^{low}		Low^{mod}
13		Low^{low}		Low^{low}	Low^{low}	Low^{low}		Low^{low}

↑↓ Arrows indicate whether the uncertainty for the 'moderate' conclusions point towards the higher or lower category.

Figure 1d The TLS is dependent on multiple complementary data sources.

In a number of the ExpGrp reports, requests have been made to further refine the data reporting around **sea lice from farms**, including details on exact dates of lice counts and fish numbers, as well as initiating the counting of sea lice in holding pens. While there may be other benefits from collecting such details, it is not clear that such refinements will make a significant difference to the overall performance of the TLS.

Estimates of lice levels derived from **sentinel cages**, **trawls** and **traps** also have their limitations which are discussed in the reports of the ExpGrp. These limitations include sampling biases of traps and trawls due to differing behaviour of infested and un-infested fish and a failure of sentinel cages to accurately reflect the level of infestation pressure across the entire width of a fjord. Again, while there may be other benefits from modifying some of these data collection methods, it is unlikely that these modifications will provide data that will substantially improve the overall performance of the TLS.

Stock Assessment Data

Information on wild salmon population size, smolt migration and smolt production are included in the TLS as described in Appendix I a[10] and b[11] in the 2020 ExpGrp report. Smolt survival and river spawning targets are central to the estimation of potential impacts on stocks at a population level. However, the database of information available for the calculation of the spawning targets for each Norwegian salmon river is limited in scope and geographic spread. This is potentially a significant data vulnerability.

The river spawning stock target is a central reference point in the TLS, the ExpGrp (2019) state:

> *Since the 2017 report, the Norwegian Scientific Advisory Committee for Atlantic Salmon has also updated its analysis of how well the coefficient of determination of the IMR's and VI's lice infestation models explain spatial variation in Norwegian salmon river's stock status (Anon. 2019). The updated models have used both the estimated harvesting potential and number of returning salmon in per cent of the spawning stock target for 2018 as response variables in regression analyses, and used data from the VI's and IMR's virtual smolts as explanatory variables from the period 2016 and 2017 (depending on the stocks' sea age distribution). We focus on the percentage of the spawning stock target when discussing these results, as this is a more logical link to the impact measurement we are interested in for the purposes of the Production Area Regulations. The analysis concludes that the attainment of spawning stock targets and impact of salmon lice (virtual smolt estimates) are important explanatory variables.*

Norway has about 450 rivers with salmon stocks. Spawning targets have been set for 439 rivers[12,13]. According to 2020 ExpGrp report Appendix 1[14], the 401 Norwegian salmon rivers with a spawning target of more than 10 kg of female salmon are included in the TLS. Spawning targets define the number of spawners that must be left in the autumn in order to reach the river's carrying capacity for juvenile salmon and is stipulated as the number of eggs (per m^2 of riverbed) or female biomass (in kg) required to utilise the river's carrying capacity and produce as many smolts as possible. The methods used to estimate stock parameters of the majority of the rivers and the figure of more than 10kg of returning females is extracted from literature that was only available in Norwegian and therefore

---

[10] Ugedal O, Fiske P, Finstad B (2020) Overview of salmon rivers. **Appendix I a** in Assessment of salmon lice-induced wild fish mortality per production area in 2020. **Report from the ExpGrp**.

[11] Ugedal O, Barlaup B, Finstad B, Skaala Ø, Sægrov H, Vollset KW (2020) **Appendix I b** in Assessment of salmon lice-induced wild fish mortality per production area in 2020. **Report from the ExpGrp**.

[12] Ugedal O, Fiske P, Finstad B (2020) Overview of salmon rivers. **Appendix I a** in Assessment of salmon lice-induced wild fish mortality per production area in 2020. **Report from the ExpGrp**.

[13] Ugedal O, Barlaup B, Finstad B, Skaala Ø, Sægrov H, Vollset KW (2020) **Appendix I b** in Assessment of salmon lice-induced wild fish mortality per production area in 2020. **Report from the ExpGrp**.

was not available to all of the EvalComm. These are important parameters underpinning the TLS and we suggest that they should be checked, and their implications reconsidered.

Based on the spawning target, a theoretical smolt production is calculated for each river (Hindar *et al*. 2007, 2019). The theoretical smolt production is then calculated on the basis of knowledge about smolt age (i.e. the number of years the juvenile salmon live in freshwater before migrating to sea as smolts) in the rivers and standard values for their survival in freshwater. A survival rate of 10% for the first year and 50% survival for each of the succeeding years is assumed. For rivers not included in Hindar *et al*. (2007), a survival estimate from the closest river is used.

The spawning target of each individual river is calculated according to Hindar *et al.* (2007, 2019). As a starting point, they use information from nine rivers where data are available for modelling of stock-recruitment (SR) relationships. Thus, spawning targets for the large majority of salmon populations/ rivers are determined from limited information on that particular river.

Recommendations:

R12 We recommend that **a systematic and substantial analysis of data collection methods and design** be carried out before modifications or additions to data sources are made, as the view of the EvalComm is that changes in site-specific sea lice estimates will likely have only minimal impacts on the overall performance of the TLS.

R13 We recommend that the ExpGrp **increase the number of rivers from which stock assessment estimates** are included in the TLS.

5. System Performance

5.1 Assessment of System Performance

The TLS has been designed to monitor and mitigate the impact of sea lice from salmon farms on wild salmonids. It is a rule-based system for capacity adjustment of salmon production, based primarily on ecological impacts. While the system has been in place now for five years, it was only in 2020 (following the 2019 round of TLS risk assessments) that there was a PA in which a 'red light' status resulted in a recommended reduction in farmed production.

Despite this low incidence of 'red light' occurrences, it should be possible to detect an effect associated with the TLS classification. Wild salmonid populations in areas consistently classified as red should show some evidence of decline, while those in areas consistently green should remain relatively stable (in the absence of any major area-wide intervention). Although, as suggested by Vollset *et al*. (2016), population-level effects of salmon lice on wild salmon cannot be estimated independently of the other factors that affect marine survival, one would expect that if the TLS was working as expected, there would be some measurable signal associated with classification. It is our understanding that as yet there has been no such analysis or review of the TLS classification, which we have referred to as the Traffic Light "systems performance". We appreciate that this is a non-trivial task, and fraught with uncertainties, but would suggest that the ExpGrp give some consideration as to how such an analysis might be undertaken.

In the first instance where a major impact on salmon survival was predicted, some of the measurable effects that could form the basis of this analysis include: spawning escapement (numbers of returning adult fish to the rivers), the juvenile stock abundance in rivers, and lice counts in sentinel cages, trawls and traps. Counts of returning fish would be of particular use here as these are the basis adopted by NASCO (North Atlantic Salmon Conservation Organization) in assessing the status of salmon stocks in individual rivers. As reported in their 2019 report on the State of North Atlantic Salmon, NASCO has assessed 2,359 rivers including many Norwegian salmon rivers. The status of stocks in these rivers is regularly assessed via ICES (International Council for the Exploration of the Sea) working groups and reported to NASCO as part of NASCO's ongoing assessment.

This database, which is available to the Norwegian competent authority, forms an independent data source not directly used in the TLS assessment process. It is the view of the EvalComm that consideration should be given as to how this might be used to evaluate and validate the outputs of the TLS. While there are potential confounding issues, such as variable off-shore mortality, it should be possible to create normalized specific return rates based on aggregated return numbers along the Norwegian coast. Over the last two assessment cycles, Production Areas 3, 4 and 5 have been red flagged as having unacceptably high impacts. In future if there is to be an assessment of the measurable 'performance' effect of implementing TLS mitigations, then this would likely begin by looking at PAs where 'red light' risk determinations and reductions in production volume occurred.

It should be stressed that the recommendation for a system performance analysis is not meant as a criticism of the system. Neither do we believe that the findings of such an analysis would necessarily be statistically significant at this stage, due to the low number of red-light PAs and the limited response data. However, as more actions and their potential effects become available, an ongoing analysis will allow the ExpGrp to fine-tune its decision-making process and generate increased confidence in the TLS as a whole. This iterative process could well be conducted within a DPSIR framework (see 5.2 below).

Recommendation:

R14 We recommend that the ExpGrp expend more of its effort and scientific reporting on possibilities for **demonstrating external validation of the approach**. (It is our impression that the focus of the ExpGrp has largely been on verifying the internal operation and predictions of the various modelling approaches.)

5.2 DPSIR: a framework for future assessment

This section presents a framework to bring together aspects of causes, consequences and responses, in an ongoing iterative assessment of the TLS.

In order to deal with complex decision-making processes involving the need to balance public good or economic requirements with the ecological and other impacts, it has long been the practice to carry out a risk analysis or cost benefit analysis. A risk analysis matrix is an efficient way to deal with a one-off impact. For on-going management of environmental impacts, a more iterative framework or process may be more appropriate and effective. These pressure-state-response systems go beyond a risk analysis to evaluate the effects of management or mitigation strategies and use this information to modify or tweak the management processes. Such processes have been used by the OECD (Organisation for Economic Co-operation and Development) and the European Environment Agency to enhance the quality of information available to experts and policy makers.

The output of the TLS is essentially a risk analysis carried out by way of a complex process of integrating several modelling, data analysis and expert elicitation steps. The key elements of the assessment process are set out in Figure 1. In general, the overall risk assessment process in the TLS follows good scientific practice. However, there is no mechanism for assessing the effectiveness of management actions nor is there provision for on-going assessment of the assumptions underlying aspects of the modelling and informing the process of expert assessment at various stages of the process. There is also scope for improved communication of scientific results in forms accessible to a broad range of stakeholders.

A number of areas have been identified where the inclusion of such an iterative pressure–state-response approach would be beneficial. These include:

- Knowledge inclusion (3.1)
- Mortality threshold estimates (3.2)
- Communication of uncertainty (3.5)
- Assessment of system performance (5.1)

The Driver-Pressure-State-Impact-Response (**DPSIR**) approach is a comprehensive framework for assessing the causes, consequences and responses to change in a complex adaptive system in a systematic way. The European Environment Agency advocates the use of the DPSIR approach (EEA Technical Report No.8/2014) as a decision-making tool when seeking to measure and evaluate the effectiveness of management and/or policy actions (Figure 2). The TLS is essentially an ecological management system and to ensure an effective balance between protection of the ecosystem and sustainable development it needs to be part of an iterative process. The process should include regular evaluations of changes in state (the status of the wild salmonid populations), the reflection of the classification in wild salmon data, the effectiveness of management actions (reduced production due to red lights) and the quality of the data and any assumptions based on it (including the mortality thresholds). The current TLS would sit well within an overall DPSIR framework. Indeed, some of the processes are already iterative, such as the modelling of sea lice infestation pressures. Many of the recommendations of the EvalComm could be addressed in the context of the DPSIR approach and would add value to the current TLS and provide a framework for its on-going improvement and fine tuning.

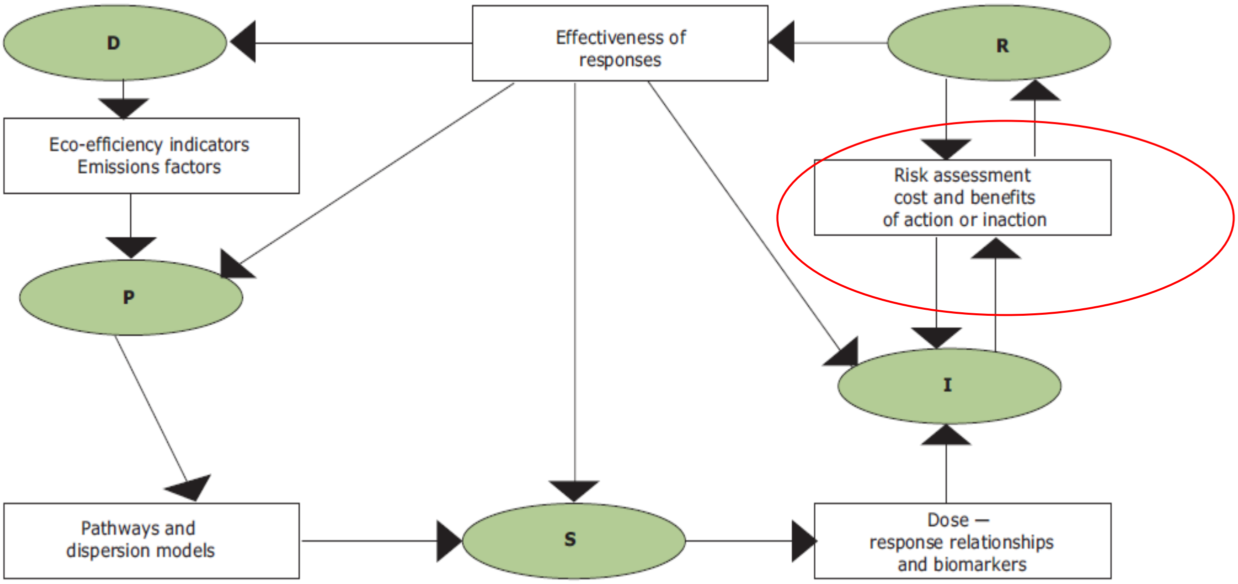


Figure 2 DPSIR (Driver/Pressure/State/Impact/Response) flowchart (after EEA): Iterative steps ensure on-going review of effectiveness

The TLS is a risk assessment matrix. It is a combination of observed data, modelling and expert elicitation. It is designed to assess the impact of sea lice infestation on wild salmonid populations. By transforming the current system into a DPSIR framework the efficacy of management actions would be explicitly assessed on each iteration (red ellipse in Figure 2) and issues related to knowledge inclusion, quantification and communication of uncertainty, mortality threshold estimates and ultimately system performance are all addressed as part of the process. The type of Drivers and Pressures which need to be regularly assessed to achieve a balanced approach have been illustrated in Figure 3a. With minor modifications to this previously outlined structure, as illustrated in Figure 3b, the drivers and pressures can align with the stated aims and goals of the TLS.

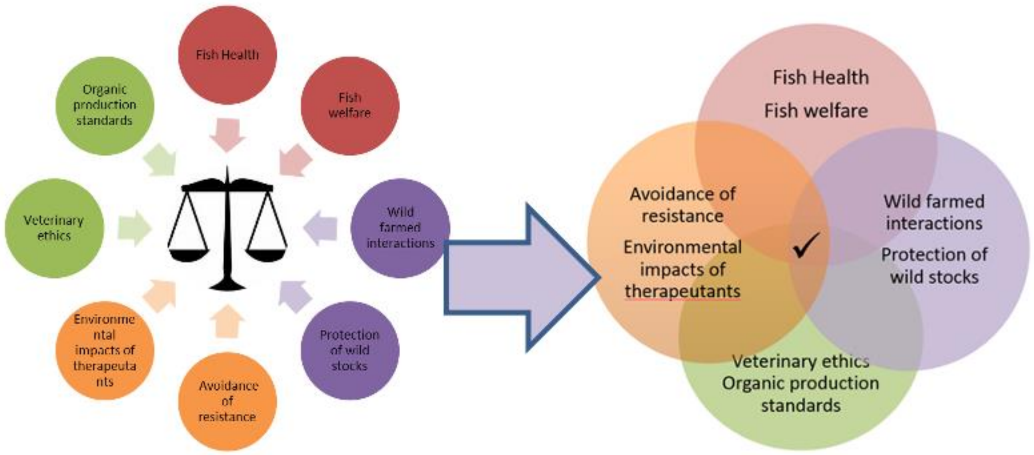


Figure 3a Regular evaluation of each variable is required to maintain the balance (from diagram in Jackson *et al*., 2017).

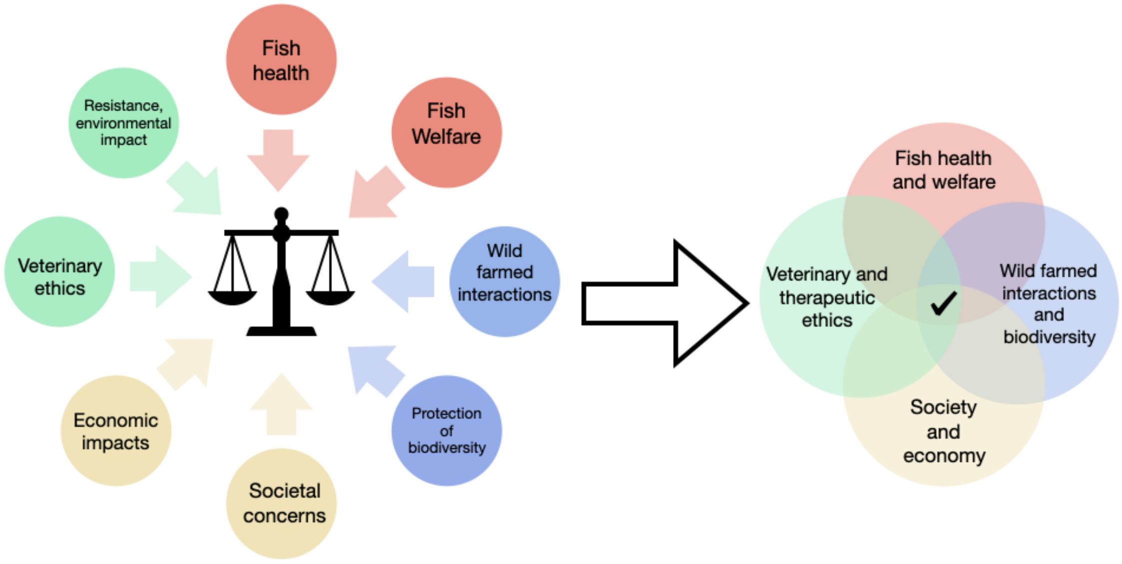


Figure 3b Modified framework, after Jackson *et al*. (2017), adding economic considerations and societal context so that drivers and pressures align with stated aims and goals of the TLS.

The EvalComm is of the opinion that the current TLS represents a comprehensive and thorough approach to carrying out a series of risk assessments using available knowledge to provide evidence-based advice to guide aquaculture policy. In this report the EvalComm has identified several key areas where improvements may be possible and has made a series of recommendations. These recommendations sit well within a DPSIR framework and would ensure the continued relevance and value of the TLS by ensuring it is constantly updated in line with new information and developments in best practice. The use of the DPSIR framework would also enhance the communication of the process and its outcomes to a wider audience of policy makers, stakeholders and the general public.

Some examples of how key areas of improvement identified by the EvalComm can benefit from the DPSIR framework are set out below.

Knowledge Inclusion (3.1)

A specific aim of the TLS is to ensure an inclusion of “all available knowledge” to arrive at an assessment of salmon lice induced mortality in wild populations. This is an important aim, and one that emphasizes broad access to and participation in scientific processes. In order to ensure there is a mechanism to implement this aim it is important that it be included as one of the indicators to be evaluated at each iteration in the DPSIR process. In line with established practices for systematic reviews or meta-analyses, it is important that a policy for inclusion and exclusion of information exists, in order to minimise bias within this part of this process.

Mortality Threshold Estimates (3.2)

In the absence of a solid empirical basis for the current Mortality Thresholds it is considered essential that their appropriateness be reassessed regularly in light of new information and as part of regular system performance reviews.

While empirical data to underpin the thresholds is not available or is incomplete it is important that a complete scientific review of available data is undertaken and documented as part of the process underpinning the formulation of a consensus expert judgement on the suitability of the Mortality Thresholds used in each iteration of the Traffic Light System assessments.

Quantification and Communication of Uncertainty (3.5)

One of the most critical aspects of translating scientific finding to policy makers and the general public is how uncertainty is communicated. Because there are numerous potential sources of ambiguity and misunderstanding, a transparent and rigorous reporting process is required. To facilitate a suitable level of feedback at each iteration in the risk assessment/costs and benefits of action or inaction loop in the DPSIR process (Figure 2), clear information around system sensitivity and uncertainty is required.

The use of more easily understood language when conveying the confidence and uncertainty associated with TLS assessments will aid communication beyond a scientific audience to policy makers, stakeholders and the general public. The ExpGrp can look to the IPCC as an example of a relatively successful protocol for such communication.

System Performance (5.1)

Two threats have been identified which have the potential to undermine confidence in the TLS, the documented shortcomings in the data and analysis underpinning the mortality thresholds, and the absence of an evaluation of the efficacy of the TLS as a management system for mitigating sea lice impacts on wild salmonids. The inclusion of a system performance assessment at each iteration of the TLS assessment using appropriate indicators in line with the DPSIR framework can provide empirical evidence of the validity and efficacy of the system. Indicators such as the level of spawning escapement, the juvenile stock abundance in target catchments, and sea lice counts in sentinel cages, trawls and traps can be utilised as a check as to whether the overall system is fit for purpose. Such regular assessments would also provide a basis for fine tuning of the methodology over time.

Recommendation:

R15 We recommend that the TLS be **framed within an iterative DPSIR process**.

6. Conclusions

In this final section of the report we provide some concluding remarks and highlight the key next steps that should follow from our assessment. We make use of the various recommendations noted throughout this report to structure these comments. A table summarising these recommendations is given below, but we address them here in a different order to which they arose in the report, in an attempt to lead the reader through a clear narrative that explains their importance.

At various points within this report, we have highlighted the important role that **uncertainty** plays within the TLS and as such we begin our comments by considering **R7** – **R10** that speak to this issue. It is the view of the EvalComm that it is equally important to identify the sources of uncertainty as they relate to the modelling frameworks [**R8**] as it is to refine the models themselves (which often appears to be the focus of the ExpGrp reports). In addition to identifying such sources of uncertainty, their implications for TLS outcomes should be explored through appropriate sensitivity analyses [**R7**]. We also note that the places within the TLS where such uncertainties exist are not always evident, which can lead to ambiguity or complete misunderstanding of the limitations inherent in some aspects of the TLS process. We therefore suggest that more transparency and rigour be applied to the process of reporting where uncertainty exists [**R9**]. It is also our view that more easily understood language should be used when conveying the confidence and uncertainty associated with TLS assessments [**R10**], as a critical aspect of translating scientific findings to a broad audience is the manner in which uncertainty is communicated. Using previously developed language models, such as those used by the IPCC, should facilitate better communication beyond a scientific audience, to policy makers, other stakeholders and the general public.

One of the areas where there remains a high degree of uncertainty is in the estimation of mortality thresholds, where two recommendations have been made. One relates to the appropriateness of the settings that are assumed to hold, which the EvalComm suggests should be reassessed on a regular basis [**R5**], in addition to exploring the sensitivity of any risk indexes to variations in threshold values. It was also recommended that *in situ* studies be carried out to better characterise the effects of sea lice infestation on wild salmonids [**R6**].

While a number of additional vulnerabilities in data extent or quality have been noted, it is the view of the EvalComm that the costs of mitigating these vulnerabilities may outweigh the benefits [**R12**]. As such, it will be important to carefully assess the realistic potential to obtain more complete data, combined with a critical assessment of the value of such data for external validation or verification of modelled outputs and policy recommendations (see also R13 and R14).

It must also be acknowledged that at times the data will be unavailable and/or insufficient to specify certain parameters, define associations, etc. In these circumstances, in addition to transparently reporting the uncertainties present (see R9 above), it may be necessary to rely on expert assessment. The inclusion of expert assessment holds not only where data are inadequate but also in the ultimate interpretation involved in transforming the various risk metrics into a final low/medium/high risk categorisation, on which the green/yellow/red classification of each PA is based.

Table of Recommendations

R1 We recommend more robust reporting of the **processes associated with knowledge inclusion** to ensure transparency and legitimacy.

R2 We recommend a **clear framework for inclusion or exclusion of sources of information** and an explicit policy on how submitted knowledge, especially reports and "grey" literature, will be evaluated.

R3 We recommend that further consideration be given to the **composition of the ExpGrp**, including the addition of expertise in areas such as scientific epistemology, knowledge inclusion, and science communication, to prompt continued reflection on such issues.

R4 We recommend that as part of the communication with stakeholders, proactive reflection be given to the **manner in which the TLS is presented**. (i.e. Is this a strictly quantitative system or more accurately a precautionary system with embedded quantitative approaches?)

R5 We recommend that the **appropriateness of the mortality thresholds be reassessed regularly** in light of new information, with careful consideration around sensitivity analysis, as part of regular system performance reviews.

R6 We recommend that studies should be undertaken and peer reviewed, to **provide data on *in situ* effects of lice infestation on wild salmon** at an individual and population level.

R7 We recommend that rather than investing effort in refining the models themselves, the models should be used to map out **sensitivity** and **identify sources of uncertainty** that can be most easily addressed with additional observations.

R8 We recommend that model, data analysis and expert elicitation results be presented with confidence intervals as much as possible to avoid a misplaced perception of their accuracy. **Proper uncertainty quantification for the models** would increase the reliability of the conclusions based on them.

R9 We recommend that a **more transparent and rigorous reporting process** around system sensitivity and uncertainty be provided.

R10 We recommend that the use of **more easily understood language** be explored when conveying the confidence and uncertainty associated with TLS assessments. A particular concern here is how these aspects of the TLS are communicated beyond a scientific audience to policy makers, other stakeholders and the general public. ExpGrp may look to the IPCC as an example of a relatively successful protocol.

R11 We recommend that the expert assessment processes within the TLS should be **described and justified in light of accepted best practice standards** that exist in the area of expert elicitation.

R12 We recommend that **a systematic and substantial analysis of data collection methods and design** be carried out before modifications or additions to data sources are made, as the view of the EvalComm is that changes in site-specific sea lice estimates will likely have only minimal impacts on the overall performance of the TLS.

R13 We recommend that the ExpGrp **increase the number of rivers from which stock assessment estimates** are included in the TLS.

R14 We recommend that the ExpGrp expend more of its effort and scientific reporting on possibilities for **demonstrating external validation of the approach**. (It is our impression that the focus of the ExpGrp has largely been on verifying the internal operation and predictions of the various modelling approaches.)

R15 We recommend that the TLS be **framed within an iterative DPSIR process**.

This should not be thought of as a ‘flaw’ in the TLS but rather an opportunity to bring issues associated with ethical research and experienced judgement into the process. However, to ensure that this is done in a credible manner, it is important that the practice and documentation of introducing expert elicitation adheres to accepted standards of best practice, as commonly adopted in related areas [**R11**]. In addition, it should be acknowledged in the final communication of results that such qualitative judgements have been purposefully included, rather than giving the appearance of an entirely quantitative and mechanistic outcome [**R4**].

Consideration of appropriate mechanisms to incorporate expert assessment leads to a number of broader issues regarding how knowledge inclusion is in general approached within the TLS. It is important that the processes by which knowledge is included in the TLS be clearly documented [**R1**]. Given that some data are taken from material that has not been formally peer reviewed, it is important that clear guidelines exist as to what will and will not be considered to be ‘valid’ sources of information [**R2**]. Due to the complexity of such processes and the implicit biases that will often exist within a group of relatively ‘technical’ specialists, it has been recommended that consideration be given to the inclusion of someone with expertise in the area of scientific epistemology among the ExpGrp [**R3**].

The EvalComm has identified a number of threats that have the potential to undermine confidence in the TLS, some of which are also noted on the ExpGrp’s reports and others of which are difficult to adequately address (e.g. inadequate data or precision is ascertaining mortality thresholds). Given these challenges, two recommendations focus on the importance of including some form of system performance assessment. Adopting such an approach can provide external validation [**R14**] of the TLS in a medium that is clear to the non-technical reader of the reports. That is to say, such a reader is unlikely to be convinced by the intricate details of the various modelling frameworks but will be able to see the logic behind clear associations between modelled classifications and the actual status of the varying environments among production areas. One particular ecosystem variable of great interest are the stock assessments in rivers within each PA. The number of rivers whose stock assessment data are currently considered within the TLS is low and as such it is recommended that the ExpGrp look to expand this source of potential external corroboration [**R13**].

The EvalComm are of the opinion that the TLS process would sit well within a DPSIR framework and that such a framework would ensure the continued relevance and value of the TLS by ensuring it is constantly updated in line with new information and developments in best practice [**R15**]. The use of the DPSIR framework would also enhance the communication of the process and its outcomes to a wider audience of policy makers, stakeholders and the general public.

References

Ådlandsvik B, Morvik A, Sandvik AD 2017. Salmon Lice Modelling at IMR - Source Term. Version 1.0 – 2017-03-29
https://docs.google.com/document/d/1zjAG6jxURC7UwVAVdZJseRl3ZDRRfZkxYc9nyXEb46A/edit?ts=5880a9e5#heading=h.bz2blmmrwuqn

European Environment Agency (2014) Digest of EEA indicators. EEA Technical Report No. 8/2014.

European Food Safety Authority (2014) "Guidance on Expert Knowledge Elicitation in Food and Feed Safety RiskAssessment," EFSA Journal 2014, 12 (6), 3734.

Hindar K, Diserud O, Fiske P, Forseth T, Jensen A J, Ugedal O, Jonsson N, Sloreid SE, Saltveit SJ, Sægrov H, Sættem LM (2007) Spawning targets for Atlantic salmon populations in Norway. NINA Report 226. 78 pages

Hindar K, Diserud O, Hedger RD, Finstad AG, Fiske P, FoldvikA, Forseth T, Forsgren E, Kvingedal E, Robertsen G, Solem Ø, Sundt-Hansen LE, Ugedal O (2019) Vurdering av metodikk for andregenerasjons gytebestandsmål for norske laksebestander. NINA Report 1303. 62 pp.

IPCC (2010) Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties, IPCC Cross-Working Group Meeting on Consistent Treatment of Uncertainties, Jasper Ridge, CA, USA, 6-7 July 2010

Jackson D, Moberg O, Stenevik, Djupevag EM, Kane F, Hareide H (2017) The drivers of sea lice management policies and how best to integrate them into a risk management strategy: An ecosystem approach to sea lice management. J Fish Dis. 1–7. https://doi.org/10.1111/jfd.12705

O'Hagan A (2019) Expert Knowledge Elicitation: Subjective but Scientific. American Statistician, 73(sup1), 69–81. https://doi.org/10.1080/00031305.2018.1518265

Stien A, Bjørn PA, Heuch PA, Elston DA (2005) Population dynamics of salmon lice *Lepeophtheirus salmonis* on Atlantic salmon and sea trout. Mar. Ecol. Prog. Ser. 290, 263–275

Taranger GL, Svåsand T, Madhun AS, Boxaspen KK (eds.) (2011) Oppdatering - Risikovurdering – miljøvirkning av norsk fiskeoppdrett, Fisken og havet, 3-2010 Havforskningsinstituttet.

Taranger GL, Svåsand T, Bjørn PA, Jansen PA, Heuch PA, Grøntvedt RN, Asplin L, Skilbrei O, Glover K, Skaala Ø, Wennevik V, Boxaspen KK (2012) Forslag til førstegenerasjons målemetode for miljøeffekt (effektindikatorer) med hensyn til genetisk påvirkning fra oppdrettslaks til villaks, og påvirkning av lakselus fra oppdrett på viltlevende laksefiskbestander. Rapport fra Havforskningsinstituttet nr. 13, Veterinærinstituttet nr. 7.

Vollset KW, Krontveit RI, Jansen P A, Finstad B, Barlaup B T, Skilbrei OT, Krkošek M, Pal Romunstad P, Aunsmo A, Jensen AJ, Dohoo I (2016) Impacts of parasites on marine survival of Atlantic salmon: a meta analysis. Fish and Fisheries 17, 714–730.