Interim Note

by the Committee tasked with

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

Evaluation Committee Members:

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

July 2021

Table of Contents

Abbreviations used in this Interim Note

- EvalCom – Evaluation Committee
- ExpGrp – Expert Group
- NASCO – North Atlantic Salmon Conservation Organisation
- PA – Production Area
- SG – Steering Group
- TLS – Traffic Light System

Introduction

This document is an Interim Note produced by the Evaluation Committee (EvalCom) 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 EvalCom was formally constituted in late 2020 with a view to produce a report by the end of November 2021. In addition to that final report, we were tasked with the delivery of a, "note with preliminary assessments and proposals for improvements", in the first instance, by July 1st 2021.

While the full set of evaluation tasks outlined in the remit for the EvalCom has begun, only initial findings are currently outlined. The purpose of this Interim Note is to give some indication as to the major foci of our work to date, which should aid in an appreciation of how the Final Report will likely be structured. Commenting on many 1,000 of pages of documentation necessitates that focus be given to a few areas of key concern, and a major goal of the EvalCom over our first six months of operation and of our interactions with the Steering Group (SG) and Expert Group (ExpGrp) has been to identify these key areas.

As such this Interim Note highlights four aspects of the TLS which have arisen as areas of focus for the EvalCom during the initial phase of evaluation. We plan to focus on these areas during the next phase of our evaluation, together with others still under initial exploration, while also providing comprehensive commentary on the set of scientific documentation produced over the past 3-4 years as part of the TLS. Before outlining these areas of focus, we would like to note that our initial assessment of the TLS is that it provides a comprehensive and thorough approach, which is world leading in terms of the attempt to link research evidence to aquaculture policy.

One challenge for the EvalCom in approaching our task was the fact that the final report and assessment produced by the ExpGrp is the result of a complex process of integrating several modelling, data analysis and expert assessment steps. No simple overview that could be used to gain a holistic grasp of this process appeared to exist and so we produced a graphic (Figure 1, that can be found at the end of this Note) to aid our understanding. In creating Figure 1 we were attempting to depict the interdependencies between data sources and analytical steps as understood by the EvalCom, 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 attempts to summarise the key elements of the assessment process. We also believe that as a reasonably complete 'system overview', aids in the identification of knowledge sources and/or knowledge gaps (see Section 1). 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, uses process-based models to answer to questions that are causal in nature (e.g. copepodid transmission) and proceeds with logical steps from empirical and theoretical knowledge towards the final assessment end points.

However, based on this overall process description, we would like to focus attention on the following separate but connected sub-processes/tasks as part of this Interim Note. *Data and knowledge sources* (Section 1) which provide the empirical evidence for the assessment. *Systems performance* (Section 2) which addresses issues of validation and questions around the forecast and predictive performance of the individual models and the process as a whole. *Mortality threshold estimates* (Section 3) which are arguably the single most important component of the assessment process, since they form the basis for the alternative intermediate impact assessments as well as in the final impact assessment. Transparent and clear q*uantification and communication of uncertainty* (Section 4) is challenging but mandatory for risk assessments in general and especially so in a process with the degree of complexity that is present in the TLS.

1. Knowledge Inclusion

In our remit, the EvalCom has been asked to, "*assess the use and choice of scientific models and methods, strengths and weaknesses, handling of risk and uncertainty, results and statistics, and quality of the assessments*," as well as, "*the transparency and verifiability in the work of the Expert and Steering groups (documentation, publications etc.)*".

It is our initial impression that the ExpGrp and SG have shown a clear and admirable ambition to include a wide range of knowledge within the Traffic Light process. 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*." This is an important aim – and one that aligns with the 2020 clarifications of the UN Human Rights Declaration's "*right to science*"[1], which emphasizes broad access to and participation in scientific processes.

In light of the ExpGrp/SG goals of: 1) analysing all available knowledge, and 2) having a committee comprised of people with expertise that position themselves to do so, we would like to highlight several points for more substantial consideration.

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 the 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, and we welcome any additional information

[1] http://docstore.ohchr.org/SelfServices/FilesHandler.ashx?enc=4slQ6QSmlBEDzFEovLCuW1a0Szab0oXTdImnsJZZVQdxONLLLJiul8wRmVtR5Kxx73i0Uz0k13FeZiqChAWHKFuBqp%2B4RaxfUzqSAfyZYAR%2Fq7sqC7AHRa48PPRRALHB

about such activities. However, since at the present time, it appears that substantial records may not be available then a recommendation that such activities are more formally recorded will likely form part of our final report.

Documentation of knowledge inclusion and open, public solicitation of knowledge – coupled with an explicit policy on how submitted knowledge will be evaluated – is important for transparency and legitimacy.

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 order to avoid bias. We would like to understand better how the ExpGrp made decisions on 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 generation. 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, what are 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? We would like to better understand the framework that was used by the ExpGrp.

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. However, 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 a wider audience. This issue, particularly with respect to dealing with uncertainty, is also considered in Section 4.2 of this report.

Future research and who carries it out?

As part of our remit, we have also been asked to consider suggestions for further research. While it is too early for us to do so in a robust way, we plan to do so with an eye toward issues of knowledge inclusion. Is there adequate access to resources for formal research in relation to various 'grey' and/or local sources of knowledge? Is there a relatively equitable distribution of scientific capabilities and ability to conduct/benefit from scientific research across various stakeholder groups/communities? We see these questions as integral to the above sections on assessing and engaging all available knowledge.

Recommendations

1.1 Although our recommendations could change with additional information, we are likely to recommend more robust reporting around the **processes associated with knowledge inclusion**, as well as an explicit policy on how submitted knowledge will be evaluated, in order to further enhance transparency and legitimacy.

1.2 Similarly, we are likely to re-comment that further consideration be given to the **composition of the ExpGrp**, including the addition of a scholar with expertise in scientific epistemology, knowledge inclusion, and evidence-based practices to prompt continued reflection on such issues.

2. Systems Performance

The TLS has been designed to monitor and mitigate the impact mediated by sea lice from salmon farms on wild salmonids. It is a rule-based system for capacity adjustment of salmon production, based primarily on environmental impacts. While the system has been in place now for 5 years, it was only in 2020 (following the 2019 round of TLS risk assessments) that there was actually a PA in which a 'red light' status resulted in a reduction. Nevertheless, it is not too early to begin thinking about how to assess outcomes in PAs that have had to reduce their capacity due to a perceived unacceptably high impact on natural salmonid populations. One would expect that if the TLS was working as expected, there would be *some* measurable effect of the actions taken. In the first instance where a major impact on salmon survival was predicted, a subsequent signal should be detectible in returning adults in subsequent years and or numbers of juveniles in the river catchment. It is our understanding that as yet there has been no such analysis or review of the effects for the TLS; 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, if for no other reason than to help them in forming their expert opinions.

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 in assessing the status of salmon stocks in individual rivers. As reported in their 2019 report on the State of North Atlantic Salmon, NASCO have assessed 2,359 rivers including many Norwegian salmon rivers. The status of stocks in these rivers is regularly assessed via ICES 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 assessment process. It is the view of the EvalCom 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. However, in the 2017 round of assessments the red light was "deactivated" and as such only very limited impact data are currently available. However, 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 were present.

It should be stressed, that the recommendation for a system performance analysis is not meant as a critique of the system. Neither do we believe that the findings of such an analysis would be statistically significant – especially with only a handful of actions and subsequent observations. Rather we recommend this analysis as a tool to guide the ExpGrp in their decisions. Further, as more actions and their potential effects became 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.

Recommendations

2.1 Explore the potential to utilise **external** data sources (such as NASCO river stock assessments) to validate the system performance around the outcome proposed by the TLS decision-making process.

2.2 It is our impression that the focus of the ExpGrp has been on *verifying* the **internal** operation and predictions of the various modelling approaches. It may be useful to expend some of the ExpGrp's time and scientific reporting on possibilities for demonstrating **external** *validation* of the approach.

3. Mortality Threshold Estimates

The estimates of salmon lice induced mortality thresholds play a significant role in all assessment endpoints (see their central location and impacts in Figure 1). Hence, the final impact assessment is expected to be very sensitive to assumption made in this part of the process. For this reason, we feel that it is important to explore the decisions around and justification of the mortality threshold estimates.

The number of lice per gram fish body mass is used to estimate the probability that an individual will die as a result of salmon lice infestation (e.g. The Expert group report 2018+2019; The Ministry of Trade, Industry and Fisheries, 27 November 2020). Based on Taranger et al. (2012) and its own assessments, the ExpGrp has used the following threshold values in its assessments of salmon lice-induced mortality among seaward-migrating salmon smolts, first-time migrant sea trout and Arctic char < 150 g (see, The Expert group report 2018+2019; The Ministry of Trade, Industry and Fisheries, 27 November 2020):

- 100% of individuals with > 0.3 lice per gram fish will die
- 50% of individuals with 0.2–0.3 lice per gram fish will die
- 20% of individuals with 0.1–0.2 lice per gram fish will die
- 0% of individuals with < 0.1 lice per gram fish will die"

The Expert Group 2018+2019 has produced an assessment of the mortality limits updated with new information 2012–2019 included (Appendix XI Assessment of the mortality limits 2019). They state, "*we concluded that we have no basis for changing the limits proposed by Taranger et al (2012). Further research is recommended into the link between lice infestation and impacts (growth, behaviour, mortality, physiology) on wild salmonids, and as the results of tank trials cannot easily be transferred to nature, further trials in nature are also recommended*."

Taranger *et al* (2012) seems to be a central document to understand in terms of how salmon lice mortality thresholds are defined and entered in the assessment; together with Taranger *et al* (2011) which is referenced in the later paper. Both Taranger *et al* (2011, 2012) and The Expert Group Report (2019) indicate that the available knowledge on infection intensity and adverse effect on wild salmon smolts is sparse and should be investigated further. The EvalCom has not been able to track down exactly where the mortality estimates thresholds originate. However, they assume that these originate from (an unpublished?) meta-analysis of the data provided in the different references cited in Taranger *et al* (2011).

Recommendations

3.1 To improve transparency of the underlying science we are likely to re-comment on the need for a **more complete scientific review** to address the provenance of the data underlying the conclusions relating to *Mortality threshold estimates of salmon lice on salmonids* as outlined in Taranger *et al* (2011).

4. Quantification and Communication of Uncertainty

The treatment of uncertainty in the traffic-light system, and scientific advice in general, has two distinct but overlapping components: Uncertainty estimation and uncertainty communication.

4.1 Sensitivity and uncertainty analysis

The final assessment produced by the expert group (ExpGrp) is a result of 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 observations that the ExpGrp reports lack a clear description of this process and as a result it is challenging to understand what data feeds into what model, or how different model results and predictions are connected to one another.

Risk assessment reports should include sensitivity analyses of the individual model components as well as on the full 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 degree either in the ExpGrp reports or scientific publications referred to in the ExpGrp reports. However, we have not yet had time to look at these in detail, so cannot make statements about their validity in general. This task will be included in the final report. The sensitivity of the final results to different sub-models and analyses is not, however, clearly described. It is not clear from the ExpGrp report how much the final assessment would change if the results of individual sub-models changed and if certain modelling choices were altered. From Figure 1, it is clear that, for example, the mortality thresholds (as lice per gram fish) play a significant role in all assessment end points and, hence, the final impact assessment would be expected to be sensitive to that part of the process.

In addition to the sensitivity analysis the ExpGrp report should also include the uncertainty analysis related to the individual model components as well as to the final assessment. In this respect the ExpGrp reports are to some degree insufficient. First, it is not fully transparent nor clear how uncertainty is treated and assessed in all individual models, and different sub-models and sub-analyses seem to use different definitions for uncertainty. 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, and (ii) questionnaires that summarize the uncertainty in and the trustworthiness of the intermediate impact assessments in each of the production areas (PA) (Appendix XII in 2019 ExpGrp report). The intermediate impact assessments include the smolt mortality indexes of IMR, SINTEF and VI as well as the sentinel cages and fish traps and nets data (see Figure 1). This process which leads to the final impact assessment is not totally 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.

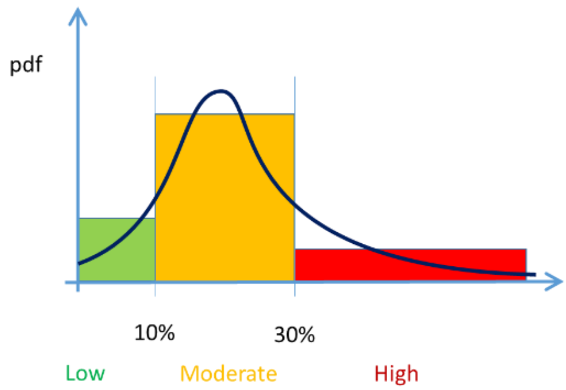
While we acknowledge that due to the high complexity of the impact assessment process, formal uncertainty quantification is hard in practice, it is doable in principle and to some extent there exist practical tools to tackle such tasks. First, the definition of uncertainty and its quantification (for example, 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. We will look to provide in our final report some more detailed advice as to how the ExpGrp might best incorporate such a treatment of uncertainty within the TLS modelling process.

4.2 The treatment and communication of uncertainty.

The approaches currently adopted within the TLS in terms of reporting uncertainty do not appear to reflect the best scientific practice, nor plain-language communication to policy makers and the general public. One of the most critical aspects of translating scientific finding to policy makers and the general public is how uncertainty is communicated. There are numerous sources of ambiguity and misunderstanding. Over the years, these aspects have been brought into focus by issues ranging from climate change to public health.

The TLS is predicated on categorizing production areas (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 system 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 opinion. The output of the procedure is to assign a category (High, Moderate, Low) to a PA and relate an uncertainty of this assignment.



The current practice focusses[2] on the uncertainty in assigning the correct traffic light category to a given PA.

The criteria[3] 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.

[2] Assessment of salmon lice-induced wild fish mortality per production area in 2019, p10.
[3] According to Memo on the description of uncertainty in the main conclusions for each production area (Nov 15 2019)

This uses the concept of the preponderance of probability, and is applied assuming that the three category assignments can to 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 3 categories is 100%, this seems to simply boil down to High, Moderate and Low uncertainty being related to the probability of the PA being in a specific category lying between 50%-65%, 65%-80% and 80%-100% respectively.

It is somewhat puzzling to understand how the two categories are expanded up into three. This process is not transparent. It 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 uncertainly 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 Low with uncertainty according to High uncertainty (50%<P2<66%), Moderate uncertainty (66%<P2<90%), Low uncertainty (90%<P2<100%)

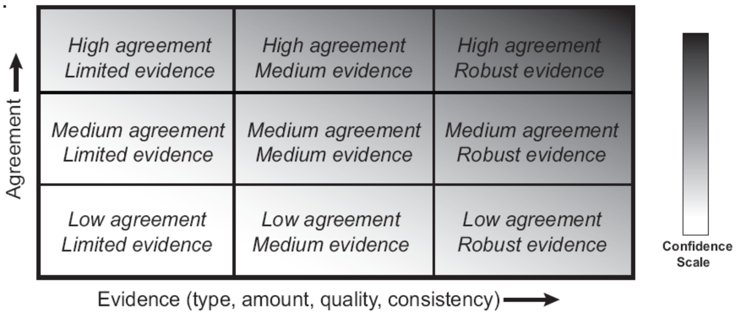
Given there are three categories, there is also a fourth uncertainty where neither P1, P2 nor P3 exceed 50%. On a 3-tier scale this is technically indeterminate. The extent to which this situation may arise in practice is unclear.

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 has over the years developed a set of protocols that have tried to standardize language regarding uncertainty. They divide this into two parts and convey the concepts of both "confidence" and "uncertainty" (or "likelihood"). Regarding this they state[4];

> The AR5 (IPPC) will rely on two metrics for communicating the degree of certainty in key findings:
>
> - **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).

[4] 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



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.

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

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

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 which is gradually gaining wider acceptance and uptake. It would be interesting to see 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

4.1 Provide a more **transparent and rigorous reporting** process around system sensitivity and uncertainty.

4.2 Explore the use of **more easily understood language** when conveying the confidence and uncertainty associated with TLS assessments. Particular concern here is how this is communicated beyond a scientific audience to policy makers, stake holders and the general public. The ExpGrp can look to the IPCC as an example of a relatively successful protocol.

References

Appendix XI Assessment of the mortality limits (2019): Karlsen, Ø.; Finstad, B. and Nilsen, F. (2019). An assessment of the mortality limits in use - updated with new information since 2012.

Ministry of Trade, Industry and Fisheries, The steering committee's assessment of salmon lice 2020 (27 November 2020). Evaluation of the status regarding salmon lice impact on wild salmonids.

Taranger, G.L., Svåsand, T., Madhun, A.S., Boxaspen, K.K. (eds.) (2011). Oppdatering - Risikovurdering – miljøvirkning av norsk fiskeoppdrett, Fisken og havet, 3-2010 Havforskningsinstituttet.

Taranger, G.L., Svåsand, T., Bjørn, P.A., Jansen, P.A., Heuch, P.A., Grøntvedt, R.N., Asplin, L., Skilbrei, O., Glover, K., Skaala, Ø., Wennevik V. and K.K. Boxaspen (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.

The Expert group report 2018+2019 (2019): Vollset, K. W.; Nilsen, F.; Ellingsen, I.; Finstad, B.; Helgesen, K.O.; Karlsen, Ø.; Sandvik, A.D.; Sægrov, H.; Ugedal, O. & Qviller, L. Assessment of salmon lice-induced wild fish mortality per production area in 2019 (2019).

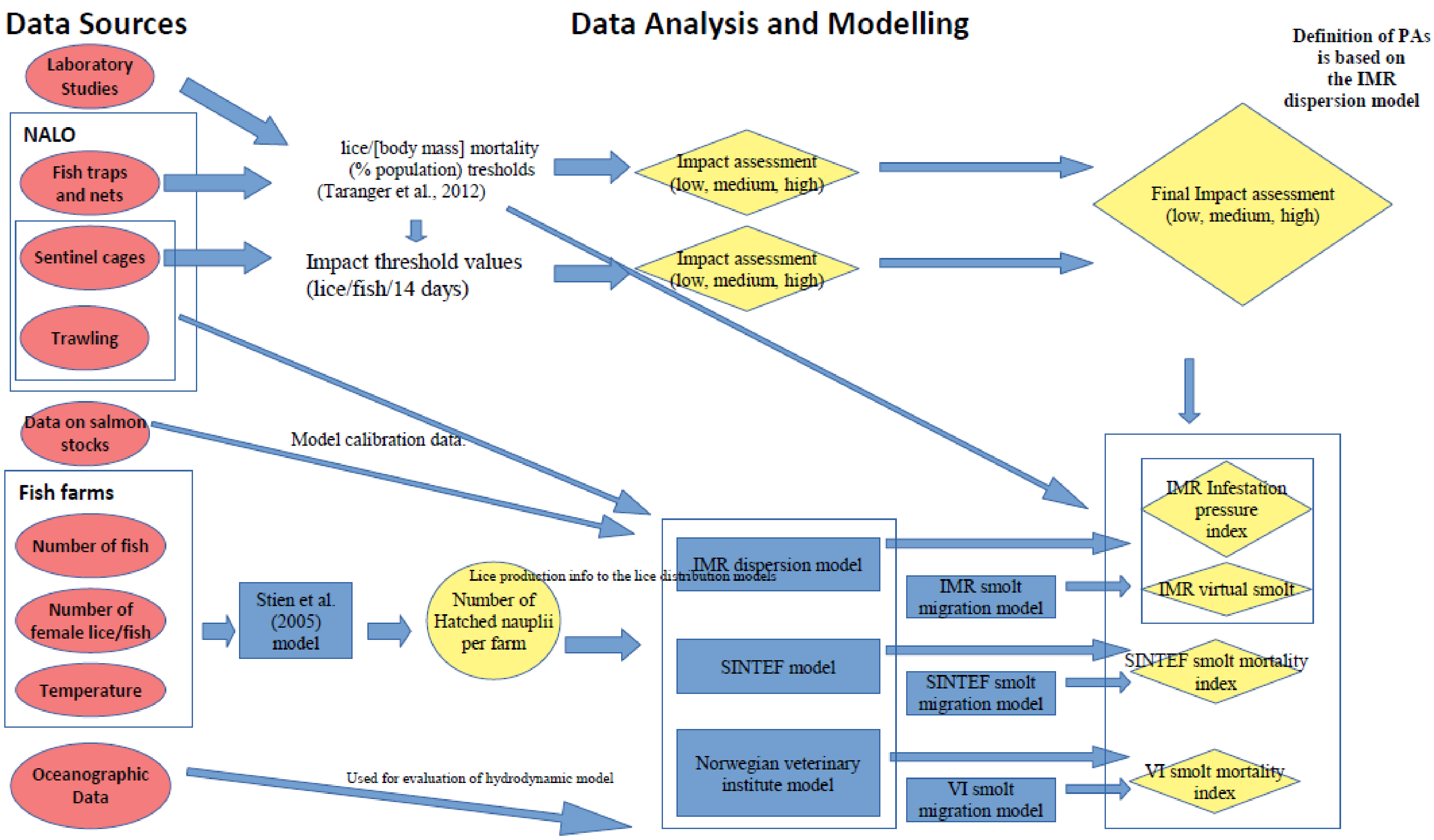


Figure 1. A flow-chart of the impact assessment in the TLS for each PA. Ovals represent data (red) or model predictions (yellow). Boxes represent models and diamonds represent assessment end-points.