The Research Council of Norway

External Evaluation of the Scientific Basis of the Assessment System for Ecological Condition in Norway

By
Katarina Eckerberg, Lena Gustafsson, Jon Moen, Andrea Nightingale,
Patrick Roose, Ullrika Sahlin and Jan Marcin Węslawski

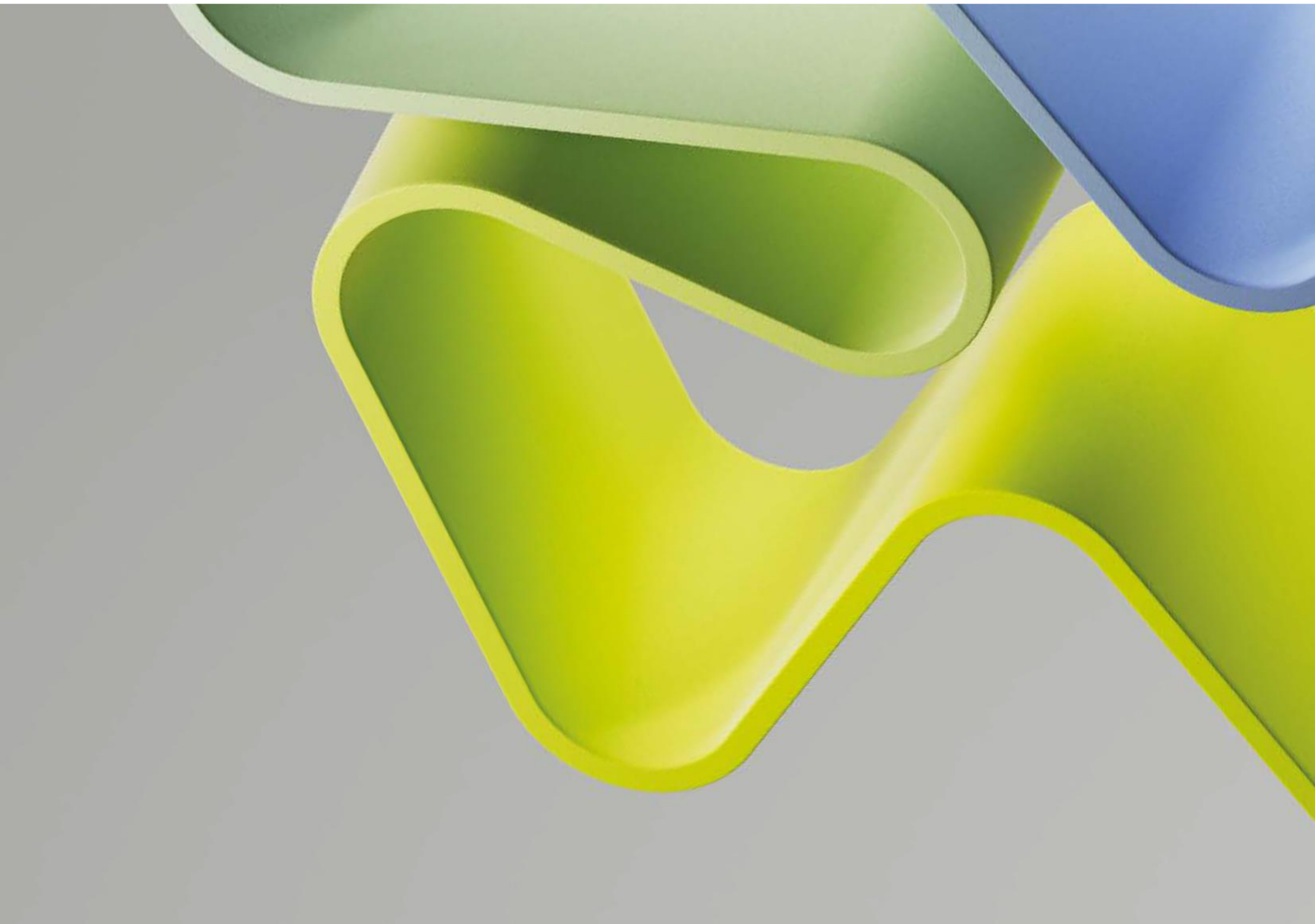

Table of Content

Boxes, figures and tables

Acronyms

CBD	The Convention for Biological Diversity
COAT	Climate-ecological Observatory for Arctic Tundra
IBECA	The Index-Based Ecological Condition Assessment
IPBES	The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem services
IPCC	The Intergovernmental Panel on Climate Change
FD	The Directorate of Fisheries
HI	The Institute for Marine Research
ICES	The International Council for the Exploration of the Sea
ITK	Indigenous Technical Knowledge
KLD	Ministry of Climate and Environment
LMD	Ministry of Agriculture and Food
MD	Norwegian Environment Agency
LBD	Norwegian Agriculture Agency
NGO	Non-Governmental Organisation
NIBIO	The Norwegian Institute for Bioeconomy Research
NINA	The Norwegian Institute for Nature Research
NPI	Norwegian Polar Institute
OSPAR	Oslo and Paris Convention to protect the marine environment of the North-East Atlantic OSPAR is so named because of the original Oslo and Paris Conventions 1972 and 1974 for protection of the northeast Atlantic Seas
RCN	The Norwegian Research Council
PAEC	The Panel-Based Assessment of Ecological Condition
SEEA EA	System of Environmental-Economic Accounting--Ecosystem Accounting
WFD	The (European) Water Framework Directive

Preface

It is a national goal that ecosystems should be in good condition and provide ecosystem services. This is stated in the Government's action plan for biodiversity. In order to manage nature holistically and sustainably, we need to know the state of biodiversity in nature, and whether ecosystems are functioning as they should. Assessment systems for ecological condition have therefore been developed in Norway.

The Ministry of Climate and Environment commissioned in 2023 the Research Council of Norway to carry out an external evaluation of the scientific basis of the Norwegian assessment system for ecological condition. The evaluation panel, consisting of seven international experts, was appointed by the Norwegian Research Council. This report represents the work, findings and advice of the external evaluation panel that was mandated to do the evaluation.

The evaluation scrutinizes and compares two assessment methods: the Panel-Based Assessment of ecological Condition (PAEC) and the Index-Based Ecological condition Assessment (IBECA). The evaluation covers the scientific basis for the two methods and comparison of their strengths and weaknesses, possible improvements and adjustments to the assessment system, including methods and scientific design. It also discusses the role of the two assessment methods in policy and public administration, highlighting any limitations of the knowledge base. The evaluation panel reviewed PAEC and IBECA reports, peer-reviewed and other articles, and conducted interviews with experts and stakeholders.

The Research Council would like to thank the evaluation panel for its excellent work. In our opinion, the panel has done a very thorough job, and we hope that the report will give valuable insight and advice both for policymakers, public administration, stakeholders, and the science community and be a common ground for fruitful discussions and improvements.

Oslo, 3.12.2024

Eva Falleth

Executive Director
The Research Council of Norway

Executive Summary

This report is an external evaluation of the scientific basis of the Norwegian assessment system for ecological condition (Fagsystem for økologisk tilstand), commissioned by the Ministry of Climate and Environment (KLD). Two assessment methods are currently used for various terrestrial and marine ecosystems: the Panel-Based Assessment of Ecological Condition (PAEC) and the Index-Based Ecological Condition Assessment (IBECA). The purpose of the evaluation is to scrutinise and compare the respective scientific strengths and weaknesses of the two assessment methods, as well as the assessments carried out since 2017, and to evaluate the suitability of the two assessment methods for evidence-based policymaking. The evaluation should cover possible improvements and adjustments to the assessment system regarding methods and scientific design, including opportunities and dilemmas, and "as far as the available information permits" discuss the role of the two assessment methods as basis for decisions in policy and public administration including any limitations of this knowledge base.

The evaluation panel has consisted of seven international experts, coordinated by the Norwegian Research Council. The panel has evaluated twelve PAEC and IBECA reports including an advisory report setting the framework conditions for the assessment system. Three peer-reviewed articles describing the respective methods have been considered. Further, the panel has interviewed five experts involved in the assessments and 14 stakeholders to better understand the methods for assessment and the potential use of the assessment results for policy and management.

The evaluation panel's overarching conclusion is that the PAEC and IBECA methods have been developed for the same purpose but in partly different institutional and policy contexts. We find that neither of the two methods follow all our proposed scientific and management criteria for an efficient assessment system for ecological condition, but that the PAEC method is better designed overall to this aim compared with the IBECA method. Since our knowledge of the processes influencing the functioning of the different ecosystems are incomplete, and indicators must be selected based on data availability and best available knowledge, we fully acknowledge the need for structured and well-documented expert judgements. In this respect, we find that the PAEC method outweighs IBECA. IBECA relies heavily on quantitative indices and aggregation of data with comparatively low transparency and high uncertainty in relation to the setting of a normative threshold of 'good ecological condition'. Inevitably, to evaluate changes over time and critical thresholds not to be exceeded will require insightful expert judgement from different perspectives and needs to be presented in a transparent manner.

We end our evaluation by proposing specific recommendations both to the PAEC and IBECA methods and to the Ministry for Climate and Environment for the future direction of the assessment system for ecological condition.

1. Introduction

1. Introduction

This report is an external evaluation of the scientific basis of the Norwegian assessment system for ecological condition (Fagsystem for økologisk tilstand), commissioned by the Ministry of Climate and Environment (KLD). Two assessment methods are currently used for various terrestrial and marine ecosystems: the Panel-Based Assessment of Ecological Condition (PAEC) and the Index-Based Ecological Condition Assessment (IBECA). The purpose of the evaluation is to scrutinise and compare the respective scientific strengths and weaknesses of the two assessment methods, as well as the assessments carried out since 2017, and to evaluate the suitability of the two assessment methods for evidence-based policymaking. Moreover, the evaluation should cover possible improvements and adjustments to the assessment system regarding methods and scientific design, including opportunities and dilemmas, and "as far as the available information permits" discuss the role of the two assessment methods as basis for decisions in policy and public administration including any limitations of this knowledge base (see Appendix A for the full mandate).
The Norwegian Research Council (RCN) has carried the administrative responsibility for the work, which implies appointing the external evaluation panel; hosting, facilitating and documenting the evaluation process; and being the contact point for the committee chair as well between the Ministry and the evaluation panel. Jonas Enge and Eli Ragna Tærum at RCN have performed the necessary support functions related to the assignment such as organising meetings, making necessary contacts and clarifying questions in the mandate and administering remunerations and travel expenses.

The evaluation panel has consisted of the following seven experts:

- *Chair* Katarina Eckerberg, Professor emerita at Dept of Political Science, Umeå University
- Lena Gustafsson, Professor emeritus at Swedish University of Agricultural Sciences in Uppsala
- Jon Moen, Professor at Dept of Ecology and Environmental Science, Umeå University
- Andrea Nightingale, Professor at Dept of Sociology and Human Geography, University of Oslo
- Patrick Roose, Operational Director at the Royal Belgian Institute of Natural Sciences, Ghent
- Ullrika Sahlin, Associate Professor at Centre for Climate and Environmental Science, Lund University
- Jan Marcin Węsławski, Managing Director and Professor at Institute of Oceanology, Polish Academy of Sciences, Sopot

The evaluation has been based on twelve reports including an advisory report from 2017 setting the framework conditions for a Technical System for Assessing Good Ecological Condition, two protocol reports introducing the PAEC and IBECA assessment methods respectively, two pilot assessments and six first assessments. Three peer-reviewed articles describing the respective methods have been considered (of which two were not mandatory readings). All materials that were mandatory readings are listed in Box 1 below. Some additional materials have been consulted when deemed appropriate by the evaluation panel and listed in the references. To guide the reader, we explain the different protocol reports and their applications to various ecosystems in Appendix B.

The evaluation panel has had two physical meetings at RCN and three Teams meetings from May to November 2024 to discuss the design of the evaluation, develop more specific questions based on the mandate, and discuss draft texts. Responsibility for this evaluation has been shared according to the evaluation panel members' respective expertise, while all members have contributed to, and stand behind, the final text.

In addition, the evaluation panel decided to conduct interviews on Teams with selected stakeholders. The interviews aimed both to expand on and validate information from the reports, but foremost to better understand the policy relevance, usefulness and limitations of the assessment methods and their applications in relation to decision making among the relevant public authorities, interest organisations and NGOs. We also interviewed five experts who have been involved in the PAEC and IBECA assessments to solicit more information both about the assessment methods as such and about their policy relevance. Chapter 5 and Appendix C contain more details about these interviews. We chose not to reveal the identity of the interviewees in this report to facilitate an open discussion. The evaluation panel expresses its gratitude to all informants for their prompt responses to our many questions and their kind support in filling in the information holes for this evaluation. We also benefitted from a final reading by one PAEC and one IBECA expert to rectify any factual misunderstandings.

In the following, we begin by describing briefly the background for developing the Norwegian assessment system for ecological condition and the PAEC and IBECA methods in Chapter 2. Here we also clarify the evaluation panel's position in relation to the core elements of the evaluation mandate to ground the further appraisal and analysis of the two methods. We then proceed to evaluating the scientific basis for each method separately in Chapters 3 and 4, while Chapter 5 is devoted to practical understanding and use of the two methods in policy and decision making based both on our interviews with stakeholders but also the evaluation panel's own scientific knowledge and judgements. **Throughout the text, we highlight the panel's evaluative judgements in bold.** In Chapter 6, we present the comparative analysis of the two methods including their strengths and weaknesses and make evaluative judgements in relation to the mandate, focusing on possibilities for improvements and adjustments of the Assessment System for Ecological Condition. Chapter 7 then summarises our overall findings in view of our mandate and presents our recommendations both for improving the two methods as such and for the Ministry of Climate and Environment to act upon.

BOX 1. Reports that were included in the original evaluation task

Arneberg, P. et al (2023a) Panel-based Assessment of Ecosystem Condition of the Norwegian Sea Pelagic Ecosystem, report 16, M-2510|2023, Institute for Marine Research.

Arneberg, P. et al (2023b) Panel-based Assessment of Ecosystem Condition of the North Sea Shelf Ecosystem, report 17, M-2509|2023, Institute for Marine Research.

Framstad, E., Kolstad, A. L., Nybø, S., Töpper, J. & Vandvik, V. (2022). The condition of forest and mountain ecosystems in Norway. Assessment by the IBECA method. NINA Report 2100. Norwegian Institute for Nature Research.

Jakobsson, S., Evju, M., Framstad, E,. Imbert, A., Lyngstad, A., Sickel, H., Sverdrup-Thygeson, A., Töpper, J.P., Vandvik, V., Velle, L.G., Aarestad, P.A. & S. Nybø (2021) Introducing the index-based ecological condition assessment framework (IBECA), Ecological Indicators 124, 107252.

Jepsen, J.U., Arneberg, P., Ims, R.A., Siwertsson, A. & Yoccoz, N.G. (2020). Panel-based Assessment of Ecosystem Condition (PAEC) – Technical protocol version 2. NINA Report 1890. Norwegian Institute for Nature Research.

Jepsen, J.U. et al (2022) Panel-based Assessment of Ecosystem Condition – a methodological pilot for four terrestrial ecosystems in Trøndelag, report 2094, M-2190|2021, Norwegian Institute for Nature Research.

Nybø & Evju (eds.) 2017. A technical system for assessing good ecological condition. Recommendations from an Expert Committee: English translation, Norwegian Institute for Nature Research.

Nybø, S., Framstad, E., Jakobsson, S., Evju, M., Lyngstad, A., Sickel, H., Sverdrup-Thygeson, A., Töpper, J., Vandvik, V., Velle, L.G. & Aarrestad, P.A. 2019. Test of the system for assessing ecological condition for terrestrial ecosystems in Trøndelag. NINA Report 1672. English translation of selected chapter. Norwegian institute for Nature Research.

Pedersen, Å.Ø. et al (2021a) Norwegian Arctic Tundra: a Panel-based Assessment of Ecosystem Condition, Report 153, Norwegian Polar Institute.

Pedersen Å.Ø. et al (2021b) Panel-based Assessment of Ecosystem Condition (PAEC) as a Knowledge Platform for Ecosystem-based Management of Norwegian Arctic Tundra, Brief Report 056, Norwegian Polar Institute.

Siwertsson, A. et al (2023) Panel-based Assessment of Ecosystem Condition of Norwegian Barents Sea Shelf Ecosystems, report 14, M-2511|2023, Institute for Marine Research.

Töpper J. & Jakobsson S. (2021) The Index-Based Ecological Condition Assessment (IBECA) - Technical protocol, version 1.0. NINA Report 1967. Norwegian Institute for Nature Research.

2. Policy Background for the Assessment System and Positioning our Evaluation

2. Policy Background for the Assessment System and Positioning our Evaluation

The purpose of this chapter is to situate this external evaluation in relation to the Norwegian and international policy context, and to set the stage for the evaluation exercise. The latter implies that we will discuss how we interpret our evaluation tasks and position the evaluation panel in view of the core concepts that are at stake in the mandate for the Norwegian assessment system for ecological condition.

Policy background

The call for measuring the status of Norway's ecosystems on land and at sea originates from the Ministry of Climate and Environment (KLD). There was a need to develop a technical system for evaluating the ecological condition as a follow-up to the Norwegian Action Plan for Biodiversity ("Nature for Life") that had been adopted by the Parliament in 2016[1]. This action plan is linked to the Convention on Biological Diversity (CBD) 2011-2020 with the so-called Aichi biodiversity targets. Hence, the Norwegian Assessment System for Ecological Condition was initiated to enable monitoring of those international as well as the "Nature for Life" goals by proposing scientific indicators and criteria to clarify what is "good ecological condition". For example, the UN and EU initiatives[2] for establishing ecological accounts (SEEA EA) use ecosystem condition as a key element. More recently, further political pressure on monitoring biodiversity is represented by the EU's newly adopted Nature Restoration Law in 2024 for both land and sea areas (which is a key element of EU's Biodiversity Strategy) aiming to restore 20% of EU's degraded ecosystems by 2030 and all by 2050. This is a somewhat lower ambition than what was agreed by the COP15 to CBD: namely to protect 30% of land and sea areas by 2030, accompanied by 23 action points (CBD 2022), which is expected to become implemented through national strategies and action plans at COP 16. In addition, the two international processes of the International Panel of Climate Change (IPCC) and the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) have served as an impetus for Norway's follow-up of climate and biodiversity policy making. In particular IPBES stresses the role of Indigenous peoples and local communities as important actors and knowledge providers in the management and governance systems for conservation and sustainable use of biodiversity.

An expert committee (with nine experts headed by NINA) was set up by KLD to report back in 2017. Its mandate was to propose a technical system that can be established for ecosystems at county/regional level (or other appropriate levels), which should be cost-effective, verifiable and reliable, and ready to use by the management authorities by 2020 (Nybø & Evju, 2017, p.10). It should be "far simpler that the system established for follow-up of the Water Framework Directive" and "be based on a limited number of indicators that reflect the structure and function of ecosystems

---

[1] Government White Paper (Stortingsmelding # 14 (2015–2016), chapter 5.3. see https://www.regjeringen.no/en/dokumenter/meld.-st.-14-20152016/id2468099/

[2] The European strategy for environmental accounts 2024-2028 and EU regulation 691/2011.

and take into account natural dynamics in ecosystem" (Nybø & Evju, 2017, p. 48). Furthermore, the expert committee should make recommendations on how often the various ecosystems should be assessed and classified, and how these assessments should build on and supplement existing relevant classification systems (see also BOX 2 below).

BOX 2. The Ministry of Climate and Environment's original design criteria for a technical system for assessing good ecological condition

- Propose scientific indicators and criteria for ecological condition in Norwegian ecosystems that at least clarify what is "good ecological condition"
- Propose a technical system that initially can be established for ecosystems at county/region level, or other professionally based, appropriate levels
- The system shall be far simpler than the system established for follow-up of the Water Framework Directive Regulations
- The focus should be on what is good condition, and not other class boundaries
- The technical system shall also be based on a limited number of indicators that reflect the structure and function of ecosystems and take into account natural dynamics in ecosystems
- In the ecosystem "open seas", the development of scientific criteria for "good ecological condition" shall be carried out as part of the work on the management plans
- The criteria for good ecological condition in ecosystem "open seas" shall be included in an overall proposal for a technical system from the Expert Committee for Ecological Condition so that a comprehensive technical system with a common framework for ecological condition can be established in all ecosystems
- Make recommendations on how often the state of the various ecosystems should be assessed and classified
- Point out which approach and degree of detail should be an aim for a technical system in the longer term

(our summary based on appendix 1, p. 48 in Nybo & Evju (2017)

Norway has implemented the EU Water Framework Directive (2000/60/EC) which establishes what features constitute "good ecological status" in inland, coastal and ground[3] waters based on a composite assessment of the quality of surface water ecosystems according to five different status classes, where the aim is to get to "good" or "high" status. The WFD also stipulates broad participation in water planning and the development of action plans by engaging relevant stakeholders at national, regional and local levels. Exceptions to achieving good or high status are made possible by identifying "heavily modified" water bodies, which are then given a different trajectory of restoration requirements. By stating that the technical system for assessing good ecological condition should be simpler that the WFD system, we note that it should not include management objectives per se, but rather "clarify what is good ecological condition, where such a condition is assessed on the basis of the deviation from a norm (intact nature)" and where managers and politicians then can set management objectives (Nybø & Evju, 2017, p.10). Since the WFD already exists, the new assessment system for ecological condition would not include the above-mentioned water bodies but all other Norwegian ecosystems (Nybø & Evju, 2017, p.48).

The expert committee's work resulted in a report which laid the ground for further assessment method development but did not recommend a specific method since the experts did not agree (Nybø & Evju, 2017). The Norwegian Environment Agency () was thereafter commissioned by the Ministry (KLD) to

[3] Groundwater is also regulated by the Groundwater Directive (2006/118/EC).

operationalise such a system for assessing the ecological condition of Norway's terrestrial and marine ecosystems. The MD allowed them to develop and test both a panel-based and an index-based method. Pilot assessments were conducted with the PAEC method for parts of the Arctic ecosystem and the Barents Sea and both the IBECA and PAEC methods were tested in the Trøndelag county for different terrestrial ecosystems. The Pilots were carried out to investigate the usefulness and reveal knowledge gaps. PAEC was thereafter applied to oceans and Arctic tundra and IBECA to forest and mountain ecosystems.

While the development of the IBECA method was headed by the Norwegian Institute for Nature Research (NINA), which is also in charge of the Nature Index (see below), they leaned upon an index-based system largely based on the classification system to establish the ecological status of water bodies in the EU Water Framework Directive. IBECA experts have taken part in the development of the UN and EU frameworks for Ecosystem Accounting (SEEA EA), which also share the notion of an index-based method for assessing ecological condition. The PAEC group included experts from the Institute for Marine Research (HI), the University of Tromsø, the Norwegian Polar Institute (NPI) and NINA in Tromsø and was inspired by the methods applied in the Climate-ecological Observatory for Arctic Tundra (COAT) programme. The terrestrial Arctic tundra assessment was headed by the NPI, while the PAEC marine assessments were headed by the HI and mandated by MD to be well connected with the existing Marine Management Planning procedures in Norway. This means that two sector agencies – the Norwegian Environment Agency and the Norwegian Directorate of Fisheries became directly involved, but several other sector agencies were also engaged in this method development process via discussions in the Fisheries Directorate Group for ecological base mapping. From our interviews we heard that the PAEC method became inspired by IPBES for further developing ecological/biodiversity aspects compared to in previous marine assessments. They found an index-based method in marine ecosystems to be difficult to apply due to the dynamics and rapid changes for instance in movement of key species that take place in marine environments together with insufficient availability of relevant data. Experts from the PAEC assessments are also involved in OSPAR (the Oslo and Paris Convention to protect the marine environment of the North-East Atlantic) and the International Council for the Exploration of the Sea assessments (ICES). Norway has not implemented the EUs Marine Strategy Framework Directive (MSFD) as such but took active part in developing the OSPAR assessment framework (including the development of indicators and their interpretation) that is also used by those OSPAR contracting parties that are EU member states for their reporting obligations towards MFSD. Norway played a major role in the latest OSPAR Quality Status Report of the Northeast Atlantic area, the result of a process quite similar to PAEC.

The Nature Index[4] from 2010, so far updated every 5 years, is developed for terrestrial, marine and limnic ecosystems and indicates the state and dynamics of biodiversity in Norwegian ecosystems by measuring about 300 indicators. It does not indicate “good ecological condition” but only deviation in the selected indicators relative to a reference value. Further, Nature in Norway (NiN) describes and classifies ecological variation across terrestrial, freshwater and marine environments since 2009, among other things, for mapping nature into various ecosystem/nature types. The mapping is (so far) not repeated for the same areas, hence with limited relevance for assessing ecological condition as such. Additional background data is found in the Red List for Species which gives an overview of developments for Norwegian species and assesses their risk of extinction, including their

[4] The Nature Index applies scaled and aggregated indicators, see https://www.nina.no/english/Biodiversity/The-Norwegian-Nature-Index

disappearing habitats and ecosystem types according to the IUCN criteria. All of these are existing information sources.

For marine environments, assessment of ecological condition has taken place since 2005 based on a Government White Paper and is incorporated into Marine Management Plans, with two expert groups - the management forum and the monitoring group - consisting of many marine stakeholder organisations[5]. Assessment processes of marine environments were thus already established at the time of setting up the 2017 expert committee and the marine management plans that include both ecological and socio-economic considerations updated every four years (the latest from 2024). As mentioned above, these marine assessments are well connected to international assessments such as those from OSPAR.

During the work with this external evaluation, an updated action plan for biodiversity was launched in late September 2024 by the two ministers of Climate and Environment and Food and Agriculture entitled “Sustainable Use and Protection of Nature” (our translation from Norwegian)[6]. It is presented as a response to IPBES COP15 and as input to COP 16. This new White Paper reiterates the need for assessing the ecological condition of Norway’s nature to fulfil Norway’s Biodiversity Goals including to further develop delivery of ecosystem services, improve management practices to this end and to protect representative areas of Norwegian nature.

Positioning this evaluation in relation to our mandate

BOX 3. Our external evaluation mandate in brief (the longer version see Appendix A)

(1) Evaluate the scientific strengths and weaknesses of the two assessment methods

(2) Evaluate the scientific strengths and weaknesses of the assessments (in practice) conducted since 2017

(3) Present possible improvements and adjustments to the assessment system, both regarding methods and scientific design, including opportunities and dilemmas

(4) As far as the available information permits, we invite the committee to discuss the role of the two assessment methods as basis for decisions in public administration and policy including any limitations of this knowledge base

The evaluation panel has tackled the mandate given by KLD (Box 3) in the following way. At our start-up meeting, a representative of MD gave further background information and developed the framework for our mandate. The efficiency of data flow and the potential for automated input of data was mentioned as well as cost-efficiency and suitability/relevance to include different types of data (see also the mandate for the 2017 expert committee). We began by studying the reports that were

[5] https://havforum.miljodirektoratet.no/en/management-plans/

[6] Meld. St. 35 (2023–2024) Melding til Stortinget Bærekraftig bruk og bevaring av natur Norsk handlingsplan for naturmangfold.

provided by the MD (see Box 1) and regarded some additional reports/articles, some of which were only available in the Norwegian language (thus with accessibility restricted to the panel members familiar with Scandinavian languages).

We followed up on details that we had not fully understood by interviewing experts involved in PAEC and IBECA as suggested by MD but decided to also interview stakeholders as described in Chapter 5. The latter interviews were not part of our mandate, but we found it necessary to provide a better picture of the policy and decision-making relevance for the assessment system as well as how the practical applications of the two methods are perceived by the relevant stakeholders. In some of these interviews we were told about other relevant reports which we also looked at. However, we are aware that other information might also exist.

Before we present our evaluation, we want to clarify how we have interpreted and delimited our mandate. We have framed our analysis around the four points in the mandate given to us by KLD (Box 3). As is stressed in this mandate we have also linked to the original design criteria for a technical system for assessing good ecological condition (Box 2; Nybo & Evju 2017), see also Appendix A ("The evaluation should be based on the 2017 advisory report").

First, we present what we regard as characteristics of an efficient system for assessment of ecological condition. For a system to deliver high quality assessments of ecological condition a number of criteria have to be fulfilled, from a theoretical ecology perspective, from a technical point of view and with regard to its implementation. We here list a number of requirements that would need to be met by a well-functioning system and will use these as an aid to make this evaluation as complete and comprehensive as possible.

From a scientific and technical point of view we stress the following:

- The technical design should be in line with international consensus regarding standards for scientific relevance and validity.
- The system should reflect important features of the structure, function and productivity of ecosystems, in accordance with important characteristics of good ecological condition.
- The selection of indicators should be based on deep knowledge on the ecological patterns and processes of the ecosystem and should be sensitive to identified hypotheses about drivers for change and supposed main natural and anthropogenic factors affecting various properties of the ecosystem.
- The identification of the state and dynamics of the ecosystem should be based on the best available scientific evidence-base.
- The system should be communicated and discussed with an international scientific audience such as through peer review.
- The procedure for the assessment should be transparent, including selection of inputs and methods for data analysis, modelling and expert judgement.
- The system should be open to revision along with new scientific knowledge.
- The data used should have suitable spatial and temporal resolution, and be updated at relevant intervals, enabling revision of the assessment.

- Uncertainty in conclusion due to limitations in knowledge should be analysed for the full assessment process.

Since our evaluation panel includes social science expertise with insights into policy and management issues, we decided to improve this information base beyond the mandatory readings. To facilitate our discussion, we state below what we believe is required for a system that is applicable to inform policy and management:

- The assessment system must be well communicated to the relevant stakeholders and audiences, and clear to all what it aims to achieve.
- The relevant stakeholders should be able to express their needs and make input into the development of the system. Which stakeholders to include depends on the purpose of the assessment system.
- Engaging Indigenous and local people who have deep knowledge of the ecosystems they use can enhance the identification of drivers and assessments of the significance of change observed; there should be possibilities for open discussions with stakeholders about which data gathering is relevant and how the method can be improved to allow for practical use of the results.
- Any judgements with concern to the result of the assessment must be scientifically well founded and explained in terms that are easily understood both by peers and by relevant stakeholders and broader audiences.
- The system should be seen as legitimate and useful to inform policy decisions by the relevant stakeholders.

It should be noted that according to the PAEC experts the mandate from the Norwegian Environment Agency (MD) was to conduct the assessments exclusively with researchers in the scientific panels, thus excluding any stakeholders directly in the assessments.

We will come back in Chapter 5 to discuss the role of the two assessment methods in policy and decision making, including their legitimacy and usefulness as perceived by various stakeholders.

The original mandate for the assessment system puts much emphasis on the establishment of what constitutes ‘good ecological condition’, which the expert committee has chosen to relate to a ‘reference condition’. Nybø & Evju (2017) refer to reference condition as “natural/untouched conditions with little pressure from human activity” (p.19) and proceed to define the reference condition in what they call a pragmatic way “as nature where the processes and structures necessary/beneficial to maintaining the diversity and functioning of indigenous species over time are safeguarded” (p.31).

But the evaluation panel wants to bring attention to the embedded social and political values in defining such a condition as well as linking it to a normative value of ‘good ecological condition’. More specifically, we agree with Hilding-Rydevik et. al. (2018) who argue that “baselines or reference conditions are always socially constructed, and thus influenced by political agendas, economic realities, preconceived ideas, and socio-cultural ‘understandings’ of human impacts on nature”. Since the aim to establish a reference condition implicitly presumes that such a condition should be strived for, an interpretation is that nature which is not in sufficiently "good condition" should be restored (as far as possible). However, the vast literature on ecological restoration emphasises the inherent social values and policy preferences when aiming for "restoring nature", including modernist, nostalgic, rational, utilitarian, ecocentric, pragmatic, socio-economic and moral restitution values in the form of community engagement (see Baker & Eckerberg (2016, p. 285-86 Table 1). We will come back to this discussion in Chapter 5 but want to point out already here that we see it as an issue to deal with for the assessments. Hence, we want to stress the importance of communicating clearly what the setting of a reference condition implies for further policy making.

We also want to clarify our position concerning the application of the so-called DPSIR framework, which was originally developed in the 1990s to structure and organise indicators in a meaningful way. We find that the concepts of drivers, pressures, state, impact, response provide a complex and non-transparent way to analyse the cause-effect relationships between social and ecological systems, and natural as well as anthropogenic factors causing change. In this report, we will use the terminology of "drivers for change" when discussing how different indicators are selected to provide science-based information about such cause-effect relationships.

Finally, we want to clarify our position regarding what constitutes qualitative and quantitative methods for assessment. We use the terms qualitative and quantitative in the following meaning: A qualitative method collects information directly from relevant actors. It can be interviews, focus groups, or a structured approach for eliciting expert judgement. The information collected can be verbal statements, as well as scores or numbers, which can be coded into quantitative information. A quantitative method collects or generates quantitative information, such as empirical measurements, statistical estimations, model predictions and remote sensing data. Both types of methods are applied within an assessment, to varying degrees. The validity of direct methods can be limited to the availability of high-level evidence. The reliability of the respective methods is based on the validity and reliability of inputs and methodology. For qualitative methods it depends for example on the selection of persons to interview or include as experts, and for quantitative methods, it depends on the type and quality of quantitative data and adequacy of the chosen methodology for quantitative analysis. Personal views on qualitative and quantitative methods can change depending on someone's experience of the methods. In some fields, direct empirical observations are seen as superior to indirect model-based predictions or expert judgement. In other contexts, formal processes for expert judgement are seen as a scientific method for assessment (O'Hagan 2019).

In addition, the application of methods can be done more or less well. For example, it is a strong characteristic to choose a statistical method considering auto correlation for inference on time-series data. It is also a strong characteristic of expert judgement when formulating well-defined questions, acknowledging the subjective nature by managing cognitive challenges and biases related to judgements, ensuring transparency and documentation of the process to elicit judgements and deal with conflicting opinions from a group of experts.

Another use of the terms qualitative and quantitative is in relation to how the outcome of an assessment is expressed. The ecological condition of a system can be expressed in a qualitative (or categorical) way, e.g. "good" or "not good". Such verbal expressions can be the result of a qualitative method, e.g. by expert judgement, or a quantitative method, e.g. a classification into categories by comparing a quantitative measure to a threshold. Since the purpose of the PAEC and IBECA methods is to assess whether an ecological system is in good condition, both will ultimately arrive at a qualitative conclusion, even though they might use different approaches to reach this conclusion. In this report we will use qualitative and quantitative to refer to methods and not the format of the conclusion.

3: Panel-Based Assessment of Ecosystem Condition (PAEC)

3: Panel-Based Assessment of Ecosystem Condition (PAEC)[7]

Overview

1. Expert panels evaluate ecosystem change.
2. Ecosystems are assessed against a set of predefined conditions ('phenomena') which are framed like scientific hypotheses.
3. Structured assessment with four phases followed by peer review.
4. Data sources are integrated to generate more comprehensive assessments than existing monitoring programmes.
5. Ecosystem change is classified into four main qualitative categories (insufficient evidence, no deviation, limited deviation and substantial deviation).
6. Stakeholders can be included in all phases of the panel's work.
7. The primary function is to assess ecological condition.

Background

The Norwegian Panel-Based Assessment of Ecosystem Condition (PAEC) is a structured protocol, following four phases: scoping, analysis, assessment and reporting, and peer-review. It is especially designed to understand change in complex ecosystems where anthropogenic and natural drivers can be difficult to isolate. Panels combine the expertise of scientists and draw on a range of scientific data, including field data, remote sensing, modelling, and biodiversity records. It is expected that panel members will have good quantitative and statistical skills, **but no specific skills in making assessments and expert judgement are mandated.**

The PAEC approach builds on methods applied in the terrestrial COAT programme[8], which has been developed for Arctic tundra monitoring by several of the PAEC experts. Monitoring in COAT applies the principle of adaptive monitoring and is based on explicit hypotheses about expected change in indicators resulting from the effects of anthropogenic and natural pressures, thus inspiring the establishment of the concept of "phenomena" in PAEC. The COAT approach involves stakeholders in the assessment protocols.

PAEC was first piloted in 2019[9] for parts of the Arctic tundra, including Svalbard and the low Arctic, and thereafter in four terrestrial ecosystems in Trøndelag. Subsequently, full assessments were carried out in the Artic tundra and three marine ecosystems: the North-Sea shelf ecosystem, the

[7] This chapter is based on: Jepsen et.al (2020); Pedersen et.al. (2021a); Pedersen et.al. (2021b); Arneberg et.al. (2023a) Arneberg et.al. (2023b); Siwertsson et.al. (2023); Jepsen et.al. (2024).

[8] Climate-ecological Observatory for Arctic Tundra: https://www.coat.no/

[9] This pilot study was published in Norwegian (NINA report 1674) and not included in our evaluation.

Norwegian Sea pelagic ecosystem and the Norwegian Barents Sea shelf ecosystem. Each of these assessments drew on results and data from monitoring programmes that are of differing time scales and less holistic in their ambitions. PAEC is intended to both consolidate these programmes and to help illustrate where data and time series gaps exist. PAEC thus has been primarily implemented in ecosystems where some large data gaps exist, although the excellent monitoring programmes allow for many indicators to have high reliability.

Operationalising the method

PAEC produces a qualitative, descriptive classification of trajectories of change in relation to a pre-defined reference condition. Baselines are established by formulating expected change in ecosystem indicators like scientific hypotheses but called 'phenomena' in PAEC. Phenomena are used to help aggregate and relate change in individual indicators to seven ecosystem characteristics (primary productivity, biomass distribution among trophic levels, functional groups within trophic levels, functionally important species and biophysical structures, landscape-ecological patterns, biological diversity and abiotic factors). These characteristics are based on similar ecosystem dimensions used by the IPBES and recommended in Nybø & Evju (2017). While not required by the PAEC protocol, the PAEC assessments have all used the same seven. While a reference state such as "undisturbed" needs to be described, it does not need to be quantifiable. The reports provide both detailed quantitative analysis of change in individual indicators, and a summary of change in the seven overall characteristics of ecosystems. Narrative descriptions of the evidence and justifications for assessment classifications are also provided. **The terminology of 'phenomenon' tends to drop out of the reports in favour of talking about indicators and their directions of change. It wasn't clear what this additional vocabulary adds other than to be a bit confusing.**

The panel can include stakeholders in all stages of the assessment, but none of them have included them to date. There is no indication in the protocol of who stakeholders might be and how local and Indigenous knowledges can inform the assessment. However, the finalised marine assessments have been discussed with stakeholders already engaged in marine monitoring groups. Stakeholder engagement in the Canadian Arctic and New Zealand offer good learning for including stakeholders in PAEC (Jepsen et al. 2024). **Our view is that engaging Indigenous and local people who have deep knowledge of the ecosystems they use would enhance the identification of phenomena and assessments of the significance of change observed.**

Defining 'good ecological condition'

The goal of the assessment is to evaluate the extent to which the current state (of indicators, characteristics and ecosystem as a whole) deviates from the reference state (here an 'intact' state). "Intact ecosystems" are those largely unimpacted by modern industrial anthropogenic activities. Good ecological condition is based on the indicators chosen to represent the ecosystem, with particular attention to capturing ecosystem structure, function and productivity. One can divide indicators into those for which *both* the value of the indicator in the reference state and in the current state can be *quantified*, and those where the current state can be quantified, but the reference state can only be described normatively or qualitatively. The latter category is by far in majority. Many indicators also have spatial dimensions such as location of the tree line as a baseline condition.

The method accommodates very different kinds of assessment of good ecological condition. For example, it can be evaluated based on a particular point in time, in relation to a 'pristine ecosystem state' or 'functionally intact ecosystem' (Jepsen et.al. 2020, p. 13) but often time-series data are used. Similarly, ecosystems can be sub-divided in a wide variety of ways, these choices often being driven by the kinds of data available. The protocol stipulates that it needs to be clearly justified how these choices are made with reference to which data sets were included and excluded. **We find that the ability to evaluate reference conditions and to divide ecosystems based on the kind of data available makes this method especially well-suited to complex ecosystems.**

Selecting and aggregating indicators

Existing and new indicators are identified in the scoping phase for which expected directions of change are proposed, ('phenomena'). Indicators are chosen based on their sensitivity to ecological disturbances, and most often have quantitative values which can be tracked over time to estimate change. PAEC utilises diverse data sources, such as biological monitoring data, satellite imagery, and long-term ecological research. If data is lacking, then qualitative reference conditions for individual indicators are defined. **We find that this multi-source approach allows for a holistic view of ecosystem conditions.**

A quantitative analysis of indicators is done in the analysis phase based on available data and from that, a systematic qualitative classification for each indicator is created. The direction and rate of change is derived from a time series analysis to assess whether a given indicator is in fact changing in the expected direction and how fast. These are scored on a scale, allowing for comparison across multiple indicators. Indicators are then aggregated by classifying change qualitatively in each indictor and combining them to assess change in phenomena. **We note that the choice of indicators and the quality of their data sets are vital and is where the greatest possibility to misrepresent important ecosystem changes lies.**

In the assessment phase, the scientific panel meets and discusses the significance and validity of the trends seen in each indicator. The indicators are brought together in a matrix which assesses the overall trajectory of change in the phenomena (see chapter 6 for an example figure). The evaluations create a matrix with validity of the phenomena along the Y axis which is classified into "low", "intermediate" or "high" based on the quality, quantity and reliability of the data. The X axis plots the assessments of deviation. The resulting matrix helps to graphically assess change in a phenomenon and demonstrates the levels of validity, reliability and uncertainty in the results. **We find the matrix approach beneficial because it provides a clear, structured, and transparent way to assess complex ecological data and compare conditions across different ecosystems.**

Evidence for deviation from the reference condition is classified into one of four categories: (1) no evidence for deviation, (2) evidence for limited deviation, (3) evidence for substantial deviation, or (4) insufficient data to assess deviation from the reference condition. This qualitative assessment aims to capture both the numerical data and the experts' interpretation of trends, risks, and ecosystem dynamics.

In this process, more weight is put on indicators for which there is a better understanding of their impact and ecosystem consequences. Classification of changes in indicators are evaluated based on how well the links between changes in the indicator, the drivers, and the consequences for the rest of the ecosystem are understood. **Uncertainties in these linkages can lead to an over- or under-estimation of change, however, the reports detail carefully how well linkages are understood.**

In the reporting and peer-review phase, the background material and conclusions from the scientific panel are written up according to the PAEC protocol. The data sources for each indicator are listed in tables and their limitations clearly indicated. A narrative accompanies the final qualitative rating, explaining how the experts interpreted the quantitative data in light of ecological knowledge and uncertainties. The assessment process often involves multiple rounds of discussion and review to refine the evaluation and ensure it reflects both the quantitative evidence and expert judgement accurately. The requirements for listing data sources and choices made for what is included and what is excluded in the analysis allow for a 'tracing back'. **The reports become a kind of knowledge clearing house and consolidation of different monitoring and data sources. This is an under-acknowledged beneficial outcome of the protocol.**

Validity, reliability and uncertainty

Validity and reliability in the PAEC method relies upon the composition of the panel, the robustness of the indicators chosen, the quality of data sets for evaluating those indicators, the discussions that ensue, the kind of statistical techniques used for analysing the data, and the qualitative evaluation of change. **This consensus-based approach decreases the potential for bias and increases the reliability of the findings. However, these approaches are well known for reducing final assessments to an evaluation that can be agreed, removing disputes and masking variation (Beck et. al. 2014). In PAEC, disputes are recorded in the reports helping to overcome these tendencies.** The marine reports discuss this problem in relation to assessment of the phenomena.

A major challenge for the assessments is how to do the statistical analysis of change in the indicators. Many indicators have varying temporal lengths, resolutions, and gaps in data. Significantly, the complexity of interactions both within indicators and between them mean that most linear analyses are inappropriate, and yet for other indicators, that is the best way to capture change. The PAEC has used normalisation and standardisation techniques to make data sets more comparable across different ecosystems and indicators, despite differences in measurement units or scales based on the respective reference conditions or thresholds. Where possible, the analysis focuses on a common time period where data are available for most indicators. This avoids biasing the assessment by comparing indicators over vastly different time frames. The terrestrial assessments used time series analysis with several different kinds of standard autoregressive models (AR), choosing the one that best modelled the data for the indicators in question. These choices are all described in the reports increasing the transparency of the method.

For the marine reports, a Bayesian approach to time series analyses was used, but deemed insufficient as non-linear trends can be highly sensitive to values at the start and end of the time series. They therefore also included augmented reality (AR) trends for comparison. **How the statistics are done thus has major implications for how and whether change from reference**

condition is captured or not. The statistical treatment of the available data may be further improved as suggested by the panels themselves by more investment in modelling of complex drivers, but the current approaches are fit for purpose.

The panel explicitly acknowledges uncertainties in their assessments by assigning uncertainty scores to specific linkages between indicators and ecosystem conditions. These scores help in understanding the degree of uncertainty associated with the relationships between specific ecological indicators. When linkages between indicators and ecosystem condition are uncertain, expert judgement plays a crucial role in interpreting data and filling knowledge gaps. **This structured approach allows decision-makers to understand where gaps exist and how much weight to place on certain conclusions.**

The reports show that significant change from the reference condition is often not detected by the method, even though the panel members have other kinds of evidence that the changes happening are important. For example, there is no detected change in the functional groups within trophic levels in the Arctic Shelf part of the Barents Sea, but the high levels of uncertainty generated by a lack of knowledge about cumulative impacts makes this finding more circumspect. Given the other kinds of changes observed, it can be expected that significant changes in this phenomenon will be more evident in the future.

The method is designed to be applied periodically, allowing for the regular updating of ecosystem condition assessments, supporting the reliability of the method over time. **While future assessments start from existing ones, the lack of long-term data inputs and consistency was of greater concern in the terrestrial ecosystems than in the marine ones where the existence of long-term monitoring programmes ensures data inputs.**

Use of experts and involvement of stakeholders

The PAEC method is based upon the use of experts, defined as those who have “acted as experts on one or more ecosystem characteristics (and their indicators and associated phenomena), and people who have acted as internal reviewers by commenting on drafts of the assessment report.” (Jepsen et.al. 2020, p. 12). The marine assessments were all led by the same person which helped ensure the reports had a standard format and emphasis. **As such, this is a method that is only as good as the experts assembled and runs the risk of being self-referential as internal reviewers become part of the expert team.**

It also includes the potential to engage stakeholders in each phase of the assessment, but in practice, this has not been done (as was requested by the MD, see Chapter 2). **Who these stakeholders should be is not defined.**

Method application in practice

Targeted ecosystems

Testing of the PAEC method was carried out for four terrestrial ecosystems in Trøndelag. Assessments have then been done in the Arctic and Sub-Arctic shelf ecosystems, limited to the Norwegian sector of the Barents Sea, by a scientific panel of 34 experts. Assessments have also been done on the Norwegian Sea pelagic ecosystem, the North Sea and Skagerrak and the Arctic Tundra including the High Tundra in Svalbard and Low Tundra on the Norwegian mainland. All these ecosystems have long-term data collection programmes but are also characterised by some significant data gaps and inconsistent time series of existing data sets.

Assessment results and interpretations

In the Arctic Tundra, efforts were made to holistically assess how anthropogenic and natural drivers of change come together. The authors made arguments for the need for more complex modelling to capture these non-linear dynamics. For example, snow cover duration (*Low Arctic tundra, Abiotic conditions*) is used to assess the effects of climate change. Both the difference between reference and current state was evaluated (a loss of approx. 30 days of snow cover), the rate of change (approx. -0.3 days/year) and also how extreme the current state. Yet, linking these findings to biotic indicators is not straightforward. The panel concluded that abiotic indicators show significant deviation from the reference condition mainly due to climate change, but that most of the biotic indicators show limited or no deviation. Therefore, they conclude that overall the ecosystem shows limited deviation from the reference condition. **We find the conclusion here to be misleading since it is acknowledged in the report that based on those indicators that do show significant deviation, it is expected that other indicators will also show significant change in the near future.**

In addition, the Tundra assessments are dominated by macro biotic species: trees, mammals, birds, with no time series data on soil microorganisms. In the reports it is pointed out that decomposers/decomposition (the most important function of micro-organisms in tundra) ought to be included as a new ecosystem characteristic in PAEC. **Since abiotic factors are likely to affect microorganisms first, with cascade effects throughout the food chain, we agree that soil biota need more attention in the future.**

In the marine ecosystems, the phenomena were defined in relation to anthropogenic drivers, specifically climate change and fishing. Interestingly, there is little mention of oil extraction in the Norwegian Sea. The reference condition in the marine ecosystems was assumed to be an ecosystem already impacted by anthropogenic activities, making an evaluation of ‘pristine’ conditions unrealistic. Thus, the indicators reflect how human activities cause change (or not). The reports indicate the need for more data on pollution and better data collection and techniques for dealing with the differences in relevant time series for microorganisms (plankton) and other species like mammals and fish.

For example, the low trophic level marine mammals reflect these times challenges (*Arctic Barents Sea, Biomass distribution among trophic levels).* For the two species included in this indicator, bowhead whales and walrus, published estimates of pre-whaling/hunting population sizes are available based on historical catch data, while data on the “current state” is supported by either a single estimate or a very short time series which does not permit time series analysis. The influence of whaling and hunting on these populations are well supported in the literature, and the very large differences observed between the qualified historical estimates and recent estimates nevertheless support the conclusion of ‘substantial deviation from the reference condition’. The assessment of the North Sea shelf ecosystem is generally based on longer time series than the other two marine assessments. In particular, time series for key biological groups (zooplankton, fish stocks) cover periods of substantial variation in the main drivers (fisheries and climate change), illustrating the significance of long time series and how conclusions about substantial deviation from the reference condition can be made with low level of uncertainty.

The reports conclude that there is substantial deviation from the reference condition in the marine ecosystems due to climate change. Some indicators show limited deviation but given the abiotic factors and functionally important species like low trophic marine mammals and seafloor habitats, they derive the overall conclusion. **A key limitation in these assessments is the location and quantity of monitoring data for different species and different parts of the ecosystems meaning that some indicators are more reliable than others.** Bringing these together to assess phenomena presents additional interpretation challenges as some indicators believed to have important influences on ecosystem change are not sufficiently valid. **The PAEC reports on marine and terrestrial Arctic ecosystems are substantially enhanced by Norway’s long-term monitoring programmes and the history of engagement with stakeholders for these ecosystems.** While data gaps certainly exist, the assessments are nevertheless able to make conclusions with a high degree of certainty. **The reports are also very well written and easy to follow, making it possible for policy makers should they wish to consult them, even if more popularised summaries would enhance their accessibility.**

Summing up strengths and weaknesses

Strengths:

- Uses existing monitoring data
- Captures complex or multi-faceted trends by drawing on the deep knowledge of the panel members.
- Assessments can be performed at different spatial and temporal scales as best fits the ecosystem and available data sources.
- Indicators do not need a quantifiable reference state to be included in the assessment.
- The method encourages periodic assessments.
- The method is designed to allow for stakeholder engagement in all stages of the work.
- The reports become an excellent source of the state of knowledge on specific ecosystems as they consolidate results from a wide variety of sources and monitoring programmes and explain all the choices made in the analysis.

Weaknesses:

- The assessment is only as good as the data sets of indicators underlying the trends evaluated. This becomes of more concern when indicators have different temporal or spatial resolutions.

- The assessments have very limited evaluation of micro-organisms in part due to a lack of existing monitoring data, yet these are likely to be changing fastest with climate change.
- Disagreements in how to interpret trends can result in assessments showing no or limited change.
- The assessments do not recognise the importance of including Indigenous knowledge.
- The reports are difficult to follow for a broader audience because of the amount of information provided.

The panel's consensus method both acts as a check on the panel members' subjective interpretations, but also can flatten disagreements in how to interpret the data**. In cases where data is lacking or confidence in causal mechanisms is low, the conclusions of the assessment often come out as 'no or limited change'. While the text is supposed to clearly signal why, these conclusions can nevertheless hide major changes that may be evident from other means.**

Our conclusion is that the assessments are only as good as the data sets underpinning them and the limitations specific data sets present for evaluating indicators in relation to wider ecosystem dynamics. The potential limitations in detecting rapid or small-scale changes could be problematic for management, especially if those changes are precursors to more significant ecological shifts. Uncertainties are carefully considered, but the protocol lacks clear guidelines on how to manage them. For example, in the marine reports the time series and sampling frames are different for different indicators, meaning that bringing them together presents additional challenges. People with skills in combining data sets from different epistemic foundations should be required as part of panel expertise.

4. Index-Based Ecological Condition Assessment (IBECA)

4. Index-Based Ecological Condition Assessment (IBECA)[10]

Overview of the method

1. Defines reference values for a good ecological condition for the ecosystem in question.
2. Determines a limit for a good ecological condition.
3. Identifies indicators for seven different ecosystem characteristics.
4. Scales the indicators to 0-1 in relation to the reference values.
5. Aggregates the indicators to the ecosystem characteristics.
6. Aggregates the indicators to ecosystem pressures or drivers.
7. Uses bootstrapping to assess uncertainty.
8. Calculates an overall index for good ecological condition.

Background

The IBECA method builds on the quantitative aggregation of scaled indicators into ecological condition indices, where the scaling relies primarily on the concept of a reference condition, representing intact nature, and limit values for what is regarded 'good ecological condition' (Töpper & Jakobsson 2021, p. 9). In this sense, it applies the same principles as the Water Framework Directive (WFD) but in a simplified way since generally only 'good' ecological status is assessed compared with five ecological classes in WFD. The method builds partly on and supplements the existing Nature Index. The scientific designers of the IBECA method claim to aim for "measurable and accessible metrics" that are "applicable with respect to informing management and decision-makers on the actual causes of ecological degradation and effect of management decisions" (Töpper & Jakobsson 2021, p. 7).

In IBECA, five main drivers as described in the Millenium Ecosystem Assessment (MEA 2005[11]) are used to relate each single indicator to ecosystem change, which also corresponds with the drivers used by IPBES[12]. IBECA aggregates ecological condition indicators into pressure indices, allowing for analysing trends by comparing index values according to available data on the pressures. It thereby aims to inform the authorities about potential management measures based on the DPSIR framework[13] (EEA 1999). The concepts of 'ecosystem condition' and 'direct drivers' in IBECA correspond to State(S) and Impact(I) (deviations from the reference condition) and Pressure(P) in the DPSIR framework (Framstad et.al 2023, p. 100).

[10] This chapter is primarily based on: Töpper & Jakobsson (2021); Nybø et.al. (2019); Framstad et.al. (2022); Jakobsson et.al. (2020); Jakobsson et.al. (2021).

[11] Millenium Ecosystem Assessment, 2005. Ecosystems and human well-being: current state and trends. Edited by Rashid Hassan, Robert Scholes, Neville Ash.

[12] IPBES webpage accessed August 2024 https://www.ipbes.net/models-drivers-biodiversity-ecosystem-change

[13] Drivers-Pressure-State-Impact-Response-Framework, see EEA 1999 https://www.eea.europa.eu/publications/TEC25

The IBECA method was tested in a pilot project for the four major terrestrial ecosystems (forest, mountain, wetland, semi-natural areas) in Trøndelag county in 2019 (Nybø et.al. 2019) and applied to forest and alpine ecosystems nationally (Framstad et.al. 2022).

Defining 'good ecological condition'

Ecological condition is evaluated via aggregation of a number of scaled indicators depending on which type of ecosystem is assessed. All indicators are scaled against a reference value. Thus, it is extremely important how the reference value is set. In some assessments, each of the seven ecosystem characteristics together with the ecosystem overall are assessed as Good (certainly good), Degraded (probably deviates from good), Very degraded (certainly deviates from good) or Uncertain.

A reference condition is called an "intact ecosystem" where historic extinctions are not considered: species introduced before 1800 CE are considered native, climatic conditions follow the normal period 1961-1990, and modern intensive or large-scale human pressures are absent. Human impact should not exceed the effect of natural pressures; humans are seen as integral in reference systems. It is of course very difficult to define the human impact for all ecosystems that should be analysed.

Jakobsson et al. (2020) present a description on how reference values may be set, and the appendices in Framstad et al. (2022) describe the reference values for all the chosen indicators. Methods range from defining reference areas to using different kind of models. Sometimes the reference values are based on expert judgement (with or without supporting models), In some cases, already aggregated indicators from the Nature Index are used, such as the "Nature index for forests" which aggregates abundance information for a large number of species and indirect measures of biodiversity. In short, a variety of methods have been used for different indicators in the assessments made. This means that the reference values that the indicators are scaled against must be seen as an "ideal" reference condition rather than as a concrete "intact" ecosystem.

According to 'good ecological condition', the ecosystem's structure, function and productivity do not differ substantially from the reference condition, defined as an intact ecosystem (see Nybø & Evju 2017, p. 26). But what does "substantially" mean? Key results in Framstad et al. (2022, e.g. Fig 3.2 and Fig 4.2) show uncertainty ranges (referred to as confidence intervals, CI) for the individual indicators, the seven ecosystem characteristics, and the overall ecological condition. The uncertainty ranges for the overall condition are quite small while the uncertainty ranges for some of the indicators are much larger and sometimes overlapping the 0.6 threshold. The uncertainty of every individual indicator is derived as the 95% CI of a distribution of 10 000 indicator values which is the result of resampling the scaled observations of an indicator with replacement and computing their average 10 000 times. The uncertainty of any aggregated index is then estimated by drawing a random value from the distribution of 10 000 indicator values for every indicator, then computing an aggregated index from the drawn indicator values and repeating this process with replacement 10 000 times. In the same technical report (Nybø & Evju 2017, p.39), the committee suggests that a value of 0.8 on a standardised scale 0-1 means a good ecological condition, but which later in Jakobsson et al. (2021) was changed to 0.6, apparently to align the method with the WFD. How different would an ecosystem with an estimated value of 0.58 be from an ecosystem with an estimated value of 0.62? **We find that there are non-transparent judgements with the IBECA method in respect to the assessment of 'good ecological condition', on what basis it is conceived, as well as how the result should be interpreted.**

Selecting and aggregating indicators

The framework for assessment is based around seven overarching ecosystem characteristics that together define and describe ecosystem structure and functioning. These seven characteristics include primary production, biomass composition across trophic levels, functional groups within trophic levels, functionally important species and biophysical structures, biodiversity, landscape patterns, and abiotic factors.

Indicators are chosen by the researcher team that does the assessment, and this is where the expert judgement comes in according to our interviews. The experts assess which indicators are desirable for the seven above-mentioned characteristics, taking into consideration what data is available preferably from Norway but sometimes also from abroad. However, the process of choosing indicators is not documented which makes it difficult to repeat the assessment without just accepting the original indicators. Most indicators come from various (long-term) monitoring programmes, which means that the data are not collected at a common scale or with a common methodology, and that representativity varies as some data are from less complete data sources. While this may seem unfortunate, it is in alignment with the original mandate to be based on existing and available scientific knowledge[14].

Some of the seven characteristics have several indicators, while others have only one or none. These may then be weighted in the final analysis, although it seems that no weights were used in the assessments of forests and mountains. In Framstad et al. (2022), the authors state that indicators may be weighted or not depending on whether they cover similar aspects of the ecosystem condition or if some indicators are assumed to reflect more important aspects of the ecosystem condition than others. **We find it unclear as to what "important" means as well as how those assumptions may be made.**

Drivers (i.e. direct anthropogenic drivers) are also identified that affect the indicators: land use/infrastructure, climate, pollution/eutrophication/acidification, exploitation/harvesting, and invasive species. The indicators may then be aggregated into one or more of these categories, and an aggregated index value for the different categories is calculated similarly to the overall index, for instance. While this may give some indications of which drivers are mostly affecting the ecosystem condition, we think that the information given by the aggregated driver indices are too coarse to facilitate management actions, not the least at low spatial scales.

In summary, there are several analyses done; indicator by indicator, aggregated for the ecosystem characteristics, and aggregated by pressure. The indicators are scaled to 0-1 and compiled to be compared with the reference condition. Finally, there is a (weighted) average of all the indicators, or alternatively, an average based on the individual indicators directly (i.e. unweighted) that gives a

[14] See Nybø & Evju (2017) Appendix 1, p. 48.

value for the ecosystem condition. Informal expert judgements may be used in this process. According to our interviews with IBECA experts, equal weight was given to each of the indicators, but as we understand it different weights can be given to the indicators depending, for instance, on the number of indicators for each average.

We are concerned that the value for good ecological condition becomes less and less transparent as aggregation and scaling remove the indices further and further from the original data. Whether or not this all makes sense, it hides a lot of ecological information behind the numbers. Even so, Jakobsson et al (2021) claim that one of the strengths with the IBECA approach is that it "provides a transparent and management-relevant quantitative approach allowing assessment of spatial-temporal variation". **We are not convinced that the aggregated values for good ecological condition, the seven ecosystem characteristics, or the pressures fulfil that claim as most of the actionable information seems to lie in the individual indicators (which are also presented).**

Validity, reliability and uncertainty

We have discussed above the somewhat unclear selection of indicators and non-transparent aggregation and scaling of those into a final index. Taken together, this implies that the reliability (and meaning) of the final result in terms of a digital value is difficult to evaluate. Moreover, the setting of the reference condition appears even more problematic in view of the discussion in Chapter 2 about the underlying value-setting for this concept. Whether the method can be repeated over time is also an issue to consider. **We find that the authors of the reports have made great efforts to specify certain aspects of the selection criteria and process of estimating a final figure, but also that much detail in this process is less evident. Therefore, to repeat the assessment seems difficult unless the same scientists are engaged in a follow-up in the future, since there are expert assumptions being made along the way.**

To what extent does IBECA distinguish variability from change over time? The overall value for the ecological condition is expressed as a proportional deviation from the reference condition which has a certain temporal aspect, but it does not indicate the rate of change unless the analysis is re-done with high enough temporal resolution. Every assessment is thus a snapshot of the condition at the time of the assessment, where time scales seem not to be discussed, and the indicator values are compared to a fixed reference value (or a fixed range of values). So far, the assessments have not been repeated so temporal changes cannot be addressed. In the interviews, it is pointed out that the whole idea of the system is to update the assessments to detect changes as discussed in Nybø & Evju (2017) but there is no clear suggestion in the IBECA reports on how often they should be repeated. However, our interviews with IBECA experts mention that method development and improved data gathering can be done intermittently while follow-assessments could follow the planned 3-year interval of updating values for some indicators in the EU's regulation on Environmental-Economic Accounting (SEEA) to which IBECA could contribute.

Spatial variability is dependent on high enough resolution in the data for spatial analyses. The assessments done so far have not been able to show spatial variation or to divide into sub-ecosystem types, e.g. for forest to distinguish between different forest types (Framstad et al. 2022).

Uncertainty in the numerical value of an individual indicator is approximated with bootstrapping (i.e. re-sampling of the original data set and deriving the indicator for every sample). For non-sample data IBECA implements either i) an elicitation process based on expert knowledge and published literature, or ii) a qualitative approximation of proportional (%) uncertainty" [like a coefficient of variation]. According to Jakobsson et al. (2021, p. 3), no uncertainty is allocated to the reference and limit for good ecological condition values as these are regarded as normative values based on best available knowledge" [cf the discussion above on socially constructed baselines]. The elicitation process would presumably depend on the expert knowledge of the assessors which would ensure that the latest scientific knowledge is included but would also make it less transparent and difficult to replicate unless the process is well documented. **We note that some applications of IBECA include information on uncertainty about the reference values and the limit for good ecological condition as a range, but this uncertainty is not considered in the rescaling, nor when sampling to generate uncertainty ranges of the indicators.** Dependencies between indicators is another source of uncertainty, which if considered, increases or decreases the value of an aggregated indicator. As a result, the uncertainty ranges for the indicators and aggregated indicators are only based on some sources of uncertainty, and the impact of the other sources is not considered, which could have a high influence on what conclusion about the ecological condition to make.

Drivers for ecological condition

The idea in IBECA is that each indicator should be responsive to pressures that relate to one of the five drivers of (1) habitat loss due to land-use/infrastructure, (2) climate change, (3) pollution, (4) exploitation/harvesting and (5) invasive/alien species. IBECA's indicators are aggregated into pressure indices according to an evaluation of which are the most critical pressures for each indicator among these five pressure categories. This grouping of indicators into pressure indices is based on expert considerations rather than empirically tested relationships.

Even if there is mentioning of various causes for environmental degradation in the background discussion of the Technical Protocol version 1.0, the IBECA method does not identify what factors could be relevant to address to improve the ecological condition, nor does it discuss change. **We find that although the method links the indicators to drivers, it is difficult to use the results of the assessments to develop proposals for policy or management measures to improve the ecological status partly due to the coarse classification of drivers**. In our interviews with the IBECA experts, they also stated that hypotheses about drivers are better discussed in PAEC than in IBECA and that this needs to be better developed in IBECA.

Use of experts and involvement of stakeholders

There is no discussion as to how the experts are chosen in the technical reports/papers. Since NINA was given the coordination task by the MD, NINA was entrusted with selecting the experts to become involved. However, the MD specifically asked for including experts from different research institutes. In our interviews with NINA experts involved in IBECA, we gathered that the recruitment of experts was not an easy task, they were either too busy or sceptical to joining due to the mandate for developing the method, in particular the task of establishing what constitutes 'good ecological condition'.

In the technical papers and in two assessments made (Nybø et al. 2019, Framstad et al. 2022), there were only ecologists involved. No managers or social scientists were involved in any of the reports. **We can understand that natural science expertise is needed to develop the method technically, but it seems that the implementation of the results has not been addressed at all.** However, it should be noted that the mandate for the expert committee does not give any guidance on how the usability of the method could be increased, even if it states that "the Expert Committee shall propose a technical system ... with an approach that is cost-effective and usable for the management authorities such that it can be taken into use in management by 2020"[15]. At the same time, the MD's instructions were to not include stakeholders in the process (see Chapter 2). Consequently, management expertise from the relevant public authorities were not consulted during the process of developing and applying the method in practice even if there were some organised reporting events that were attended by the most relevant national agencies along the process.

Method application in practice

Targeted ecosystems

The IBECA method has been tested in a pilot study on regional level in Trøndelag county and implemented on a national level for forests and mountains. The spatial coverage possible to evaluate is directly dependent on the availability, quality and resolution of the data. Some ecosystems will always have attracted more interest than others and as such have more long-term monitoring data. This includes economically interesting ecosystems (such as forests) or ecosystems that are subject to specific interest, management or law (such as the WFD). Others will have gone more "under the radar" and have less documented information.

For instance, the assessment on semi-natural grasslands concluded that they did not have enough data for a full analysis. In the forest assessment, they did not subdivide the main ecosystems due to a lack of indicators that could be assigned consistently to different subtypes. Although the municipality level was not part of the mandate, an observation was that there were not enough data points in the National Forest Inventory to represent ecological condition at this spatial level. They calculated condition values for broad regions (North, Central, West, East and South) but only values for the national level are in the English version of the report presented as "the regional condition values are rather similar" (Framstad et al. 2022, p. 24). **Publication in English would increase transferability of results to an international audience.**

Assessment results and interpretations

The national assessment of forests showed an aggregated index value clearly below the limit for good ecological condition (0.42), including low values for large carnivores, coarse woody debris, dead wood

[15] See Nybø & Evju 2017, p. 48, Appendix 1.

total, deciduous trees, areas without infrastructures, and old forests. This is something that is well-known for managed forests in Fennoscandia (and perhaps elsewhere).

The national assessment of mountains showed a value just above the limit for good ecological condition (0.68) but with low values for Arctic fox populations, small rodents, and wolverine populations. The functional characteristics of the ecosystem were also found to be low. The overall index is above the limit because of rather high indicator values for biodiversity, landscape ecological patterns, primary production and abiotic conditions. So, the ecosystem is in good condition in terms of basic structures and function (i.e. relatively low levels of human disturbances and infrastructure), but key animal populations are doing badly. What should we then conclude about the ecosystem condition? That the aggregated index says that it is in a good ecological condition, or that that the ecosystem is doing poorly because the trophic and functional structure of the ecosystem is disturbed? We also know that arctic fox populations are closely related to small rodent cycles. How does the use of dependent indicators affect the result of the assessment? The assessment thus, as for other ecoystems, opens up the possibility to target not only the aggregated index value but also to put emphasis on individual indicators, and to the seven ecological characteristics.

The Trøndelag assessments of wetlands and semi-natural land are inconclusive because of lack of suitable indicators (Nybø et al. 2019). The authors conclude that monitoring should establish a better knowledge base. **We suspect that this lack of indicators will be the case for many ecosystems, and that significant effort would be needed to rectify this with improved ecological data gathering.**

Presentation of results

The report from NINA on a national assessment of forests and mountains together with the NINA report on wetlands and semi-natural land in Trøndelag comprise about 500 pages together. We understand that there is a need for pilot-projects to document what they do, but this is clearly not useful for management. **We even found the texts quite difficult to grasp even for academic researchers like ourselves.** How would a report directed towards management look like? Which data or results are relevant to include? We gather from our interviews with scientists as well as stakeholders that there have been some webinars arranged to spread the information, which is excellent. Also, we understand that there were no specific resources allocated to wider communication other than within and by the national agencies. Our interviews with the IBECA experts reveal that it is unclear whether their research institutes should be responsible for communication with, for instance, municipalities and regions as well as industry. **We also note that even some of the most relevant stakeholders have had difficulty in grasping the message that the IBECA assessments have produced, which suggests a need for guidance of how to conclude based on the aggregated indicators and communication improvement.** In Chapter 5 we will discuss further how the results have been communicated and understood by various stakeholders.

Summing up strengths and weaknesses

Strengths

- Uses existing monitoring data.

- Relatively easy to repeat the assessment, or to add new data (as long as the original expert judgements are accepted).
- Aligns well with other systems and tasks, such as the Nature Index and the UN framework for ecosystem accounting (SEEA UN).

Weaknesses

- Non-transparent and undocumented expert judgements, especially at the start of the process.
- Aggregation of indicators and indices in several steps which hides ecological information and makes it difficult to interpret the assessment.
- Connections between indicators and drivers are not well worked out.
- Quantitative uncertainty ranges do not consider uncertainty in reference values or dependences between indicators.
- Guidance on what conclusion to make on ecological condition based on aggregated indicators is not always clear.
- Difficult for stakeholders to understand (or sometimes accept) the overall assessment which makes it difficult to translate to management decisions.

An advantage with IBECA is that it resonates with the WFD in respect to using the concept of ‘good ecological condition’ against a reference value, which is a well-known concept among key stakeholders. As mentioned in Chapter 2, the WFD process involves a large number of stakeholders at all administrative levels to follow-up with management and action plans. However, there was no explicit objective in the MD’s mandate to IBECA to develop such plans, but this must be done in other processes by other means. We will come back to this discussion in chapters 5 and 7.

We find that the IBECA-method is less transparent than perhaps intended. The definition of an “ideal” reference condition, choices of indicators, and scaling and aggregation processes all require a substantial number of assumptions and expert assessments. This also makes the method less standardised than suggested, and while the overall value on the ecological condition can be compared between ecosystems, it is not clear what you are actually comparing if different indicators or reference conditions are chosen for the different ecosystems. We are also not convinced that the method gives information of high management relevance.

5. The Role of the Two Assessment Methods in Policy and Public Administration

5. The Role of the Two Assessment Methods in Policy and Public Administration

Our mandate stipulates that we should also evaluate the role of the two assessment methods as a basis for decisions in public administration and policy including any limitations of this knowledge base. To enable such a discussion and evaluation, we found it necessary to interview representatives from the relevant public administrations and interest organisations. Eleven interviews were carried out as a semi-structured conversation with a total of fourteen individuals. Three additional interviews were held with five PAEC and IBECA experts that partly covered questions relating to the policy application of the methods. Interview questions and represented organisations are listed in Appendix C. The purpose was to unfold different perspectives on the two assessment methods and their relevance to policy work rather than a strict question-and-answer exercise. We decided to make these informants anonymous to allow for more personal reflections as they were interviewed in their personal capacity rather than as a formal representative for their respective organisation. In the following we present what was said concerning PAEC and IBECA respectively under each subheading (mostly based on our questions to stakeholders, see Appendix C). The text also builds on documentation found in different reports.

Knowledge and legitimacy of the methods

Our informants were selected on the basis that they were key individuals in different organisations who are responsible for the application of the assessment system for ecological condition. Not surprisingly, all were therefore well informed especially as to how the assessment system for ecological condition more generally is understood and works for practitioners. **While most of the interviewees expressed that they had not studied the respective methods in any detail, the majority still said that they trusted what the scientists had done and accepted the resulting reports about the ecological condition**.

However, there was a distinct difference between PAEC and IBECA in respect to legitimacy of the method considering how PAEC was applied to marine and IBECA to forest ecosystems. We found that this difference is related to the legacy of applying the assessment method in the different ecosystem types. Although Arctic tundra was one of the ecosystems in which PAEC was first applied, the main application has been in marine ecosystems. Marine assessments have been run as part of the marine planning process for about 20 years (the first marine management plan was produced in 2006), where scientists and managers have collaborated for many years in the two marine management groups (the Monitoring Group and the Management Forum). In this process, the PAEC method allows for stakeholders to have input into the process of interpreting the results as coordinated by the Monitoring Group and led by the Institute for Marine Research (HI). However, we were informed by stakeholders that the assessments done through the PAEC method were exclusively involving scientific experts in practice, even though the PAEC protocol allows for bringing in other types of knowledges about monitoring needs. Hence, stakeholders only become involved when discussing the PAEC results. We asked specifically whether Indigenous and local knowledge is represented in this process and were told that this is not the case. Our question seemed to be surprising, although the IPBES stresses the importance of including such knowledge in the management and governance of biological diversity, where monitoring is key.

The results of the Monitoring Group are thereafter discussed among stakeholders in the Management Forum and reported back to Parliament. Both the Monitoring Group and the Management Forum consist solely of the relevant public agencies (see https://havforum.miljodirektoratet.no/om-faglig-forum/organisering/) but the proposals for marine management are published and inputs from various stakeholders encouraged (see https://havforum.miljodirektoratet.no/involvering/innspill-til-faglig-grunnlag/). We note that while the earlier marine management plans were mostly commented upon by marine industry, more lately environmental interest organisations have also made their inputs (albeit we found no indication of indigenous/local knowledge contributions in those consultation processes). Due to this rather inclusive process, **we found that the PAEC method as applied to marine ecosystems is well established and trusted among all those stakeholders we interviewed.** None of our interviewees brought forward the PAEC Arctic tundra assessment, which implies that we have not analysed stakeholder viewpoints on the legitimacy for this ecosystem assessment. We also learned that for marine ecosystems, HI initially accepted the data collected in the Nature Index as a basis for assessment, but that this changed around 2015 when it was revealed that albeit notable change in marine ecosystems, the updating of the Nature Index from 1990 to 2015 did not show such change. **Hence, HI concluded that an index method would not detect ecological change in marine ecosystems.**

By contrast, the IBECA method was developed in 2019-22 by scientists alone without any direct connection to ongoing natural resource policy and management processes. While stakeholders from various environmental organisations convey great faith in how the scientists have gone about to identify 'good ecological condition', **the IBECA method – and its results - are questioned particularly by those stakeholders representing public/private forest management**. For instance, the published results of the Norwegian version of the Assessment of Norwegian Forests and Mountains in 2020 became criticised and debated in public media. As far as our interviews disclose, the IBECA method is however not disputed for mountain areas, most likely due to a better score for the ecological condition for mountains compared to forests combined with a different policy context related to its management. Our interviews revealed that the application of a 'reference condition' for Norwegian forests is highly contested on the grounds that such a value appears to be rather arbitrary, and even that it might originate from a biased view on forest's ecological condition since Norwegian forests have a very long history of being managed. Yet, one of our informants noted that marine areas have also been impacted for a long time, restricting possibilities to identify an 'intact state' or 'reference value'. It was also unveiled in the interviews with forest sector stakeholders that while the legitimacy for using data from the National Forest Inventory is high, there is much less faith in the ecological data from NINA. Some informants expressed that some of those indicators have little to do with forest management practices (e.g. for ungulates and predators) but are affected by other policies. Also, it was mentioned that some IBECA indicators overlap so the aggregation of the final index figure might imply 'double counting' (such as the amount of deadwood, the presence of ungulates and predators which are both separate indicators and present within the Nature Index). There were hopes expressed both by one interviewee from environmental interests and another from the forest sector that the inclusion of more indicators – particularly from the National Forest Inventory - would help to improve the relevance and legitimacy of the IBECA method, hence in sharp contrast to the original mandate given to the expert committee to create a simple system. **Another general observation**

from our interviewees was that while the PAEC method was deemed explicit about its use and justification of expert judgements, the IBECA method was regarded as far less transparent in how such judgements are being made even if they are highly present.

As a way to mediate the low acceptance of the IBECA assessment results in forests the two ministries Climate and Environment and Agriculture and Food jointly commissioned the two agencies MD and LBD respectively in November 2022 to investigate whether a common understanding could be reached on how to assess ecological condition for forests (see Miljødirektoratet & Landbruksdirektoratet 2023, p. 29). Additional sector agencies also took part. Interestingly, this can also be viewed as a criticism against the original MD mandate given to both the PAEC and IBECA methods (see Chapter 2). This task was added to their previous task from September 2022 to investigate different policy measures that would be relevant to maintain/improve the ecological condition in forest (the latter task also included several other agencies). They agreed on 13 properties to assess 'good ecological condition' but without specifying a reference condition and what should be considered to count as 'good'. The criticism concerns both the data availability for applying the index method to forest ecosystems and the perceived goal of reaching 'intact state' in all forests. Their report[16] was sent for public hearing where 19 organisations made official inputs.[17] For example, the Norwegian Institute of Bioeconomy Research (NIBIO) was highly critical to keeping the figure 0.6 for 'good ecological condition' of forests as a 'scientific threshold set by natural science'. They pointed at the need for clarifying what role the 0.6 threshold will have in future work, and that further scientific assessments are needed for different forest types and regions (including production forests, non-productive and protected forests) to lay the ground for management goal setting. This critique became published in Norwegian news 'Nationen'. There is hope, according to MD, that the scientific assessments would help to create consensus between different stakeholders over the ecological condition of Norwegian nature. As for the apparent tensions in the forest sector, we were informed that the dialogue between sector interests continues.

Still, **to become better informed about what constitutes "good ecological condition" in various environments including forests is seen as important by all stakeholders**. **But there is less consensus among stakeholders as to how to interpret and use the results for developing action.** Hence, we find in our interviews that most stakeholders put great emphasis on the applicability of the results from the assessment in policy and management practice to guide future management planning and development of policy measures, in addition to their use in national and international reporting systems. We further note that there is considerable difference in how the two coordinating scientific institutes link to the most economically powerful stakeholders: the HI is working closely with the marine sector, but NINA much less so with the forest sector, which might also partly explain the different patterns of trust between scientists and stakeholders in relation to the PAEC and IBECA methods.

Hence, we conclude that the difference between the PAEC and IBECA methods is notable in view of how the respective assessment processes relate to ecosystem management policy and decision-making processes, where PAEC - when applied to marine ecosystems - is clearly connected and IBECA explicitly separated. This difference between the two methods clearly plays

[16] MD and LBD, 2023. Kunnskapsgrunnlag om økologisk tilstand i norsk skog og utredning av tiltak, Rapport M-2597.

[17] See https://www.regjeringen.no/no/dokumenter/invitasjon-til-a-gi-innspill-til-rapporten-kunnskapsgrunnlag-om-okologisk-tilstand-i-norsk-skog-og-utredning-av-tiltak/id3016317/

out in how stakeholders view the legitimacy of the assessment methods and their application in practice where the IBECA method is contested among certain stakeholders linked to economic forest resource management, but PAEC appears not to be so (despite the presence of economic actors in marine systems exploitation). However, we found that the interviewed representatives of environmental organisations all expressed trust in the IBECA method. We notice that this difference between the two methods can be explained by the level of trust between (some) stakeholders and the scientific experts carrying out the assessments. Several of our interviewees pointed out that the longer tradition of environmental assessment in marine environments and the associated policy processes connected to stakeholders might explain the higher trust in the PAEC method. **The evaluation panel finds that resolving the disputes about the relevance of the IBECA method will require considerable effort and inclusion of the relevant interests and knowledge to align the expectations from the various stakeholders.**

Views on the existence of two parallel methods and their presentation

None of our interviewees thought that it is a problem that two parallel methods exist for marine and forest and mountain environments. Several pointed to the longer legacy of the marine management plans (with regular assessments), but also to the dynamics of marine environments with more rapid change and movements of key species compared to terrestrial milieus (that prevents the establishment of an ‘intact state’ and ‘reference value’ as such). In the best of worlds, some of our informants said, we should have a common and standardised system for all environments and an index method is generally easier to understand for laymen: more transparent, likely to be easier to repeat, and more useful for international reporting and development of natural capital accounting. However, the current data availability implies that a common method for all ecosystems is unrealistic. And as one informant stated, the general acceptance of the method might be more important than its consistency across ecosystems.

All stakeholders agreed that **the presentation of results is key, but also that the reports so far have been too thick and detailed for many to be able to read** (however, one of them thought that the IBECA method could have been even better justified as to how the scientific judgements were made). Discussions on the systems tend to become very technical in the reports and hard to follow for a policy person. The summary is most read, but it was also mentioned that webinars had been held by NINA to explain the result of the IBECA assessments, which were well appreciated and could be done at greater extent to communicate the results more widely.

The evaluation panel concludes that the current communication outlets for the assessments carried out so far are not well designed for the purpose of reaching out to the relevant stakeholders. We find this problem to apply to both assessment methods, but that it is even more pronounced for IBECA due to excessively technical language being used in its reports along with non-transparent evaluative decisions about the ecological condition.

Usefulness for policy and management

An issue that several informants expressed was that **the main purpose with the assessment system for ecological condition should have been better explained from the onset, as there are diverging interpretations in this respect**. Some of our interviewees (notably those from national authorities) told us that it is merely to assess the ecological condition in order to report this against political targets at national/international level, while others emphasise that the results should be used to measure change over time as well as to develop restoration measures for different regions/ecosystem or management types/habitats. Several informants expressed a strong wish for the assessment results to be useful to inform management, also for counties (fylker) and municipalities. Particularly those stakeholders who expect the assessments to be input to management planning and concrete action at lower levels of governance seem disappointed. **Expectations of the use of the results thus vary considerably among the relevant stakeholders, even if many accept the report results as scientifically valid**.

There was widespread agreement across stakeholders that action plans are needed as follow up of the assessments since these are geared toward knowing which ecological factors are in "good status" and which need strengthening. According to an MD representative, the assessments should create a common understanding with a scientific base/knowledge platform which directly spurs action. Still, we also heard from national agencies who argued that without agreement on the ecological condition of forests it will be impossible to develop action plans and specific policy measures. One such national level representative expressed that acceptance might be more important than scientific perfection. **We found that while the PAEC method can function in this way via the marine planning groups, the IBECA method applied to Norwegian forests is too contested to form a basis for further action**.

Moreover, while some see the setting of a 'reference value' as a purely scientific exercise, several stakeholders emphasised that it implies that such a condition is implicitly preferred and therefore requires action, but that the assessment system as a whole is currently not designed to this end. A main problem that many stakeholders mentioned is that the assessments are too coarse as they are done for large areas rather than for regional/local circumstances. This in turn means that developing proposals for policy measures renders difficult, as ecological conditions vary greatly between different areas and subtypes of ecosystems such as different forest types or particular fjords. For instance, it was argued that regional/local assessments should be developed for specific fjords in need for ecological restoration measures (which could be covered by the WFD), and at local level for municipalities to use them in their land use planning (and in natural capital accounting). There are current desires from municipalities to develop goals for natural capital accounting within their territory, where assessments of ecological condition could provide important input if aligned with local ecological data information. The MD mentioned that municipalities will soon be required to carry out landscape planning, where the ecological assessments could be relevant for modelling also at local level. Our interviews with KLD as well as with IBECA experts suggest that several municipalities are asking them for such information, which is (yet) not readily available. In addition, several stakeholders stress that the assessments should be conducted on a regular basis to detect human impact on the environment as well as impacts from climate change.

One informant stressed that a problem with both methods is their unclear link to pressure factors/drivers, which makes it difficult to identify relevant mitigation actions. Developing a possibility to measure effects of different management practices would thus be desired to increase the usefulness of the method, but this might perhaps require a different (parallel) assessment system. Informants at KLD as well as from the forest sector stressed that for effective use of the assessments the more detailed judgements and information on indicators and ecological characteristics must be used rather than the final aggregated results. However, KLD also reiterated the need to simplify the assessment processes in comparison with the WFD which requires a lot of resources.

Our interviews showed that both assessment methods are used for policy development at national level, namely for national level marine management planning as well as for the national strategy for wetlands, while a forest action plan is yet to be agreed upon. According to MD as well as IBECA experts, data provided by the IBECA method can be used for natural capital accounting, since it provides data that resonates with international standards. As earlier mentioned, the PAEC assessments for marine ecosystems are relevant for international reporting to OSPAR and ICES about the marine environments. It should be noted that while the PAEC method assesses the ecological condition as such, socio-economic factors are also acknowledged when developing the marine management plans. Our interviews with stakeholders engaged in the marine management planning processes indicate that conflicts such as with the fishery and oil industries are recurrent when balancing ecological goals with socio-economic ones when trying to translate the management plans into action. This implies that the PAEC results for marine ecosystems are not directly feeding into action but are rather used as one of several inputs to subsequent political processes.

The KLD uses information from the assessments for budgeting the need of resources for protecting biodiversity/nature. The new White Paper in 2024 linked to CBD (Meld 35: 2023-24) stresses this need and might lead to new political initiatives to this end in the near future. Improving the relevance of ecological data as well as ensuring long-time series to detect environmental change will be key in respect both to being able to conduct regional and local assessments, to provide input into natural capital accounting and to develop management and action plans within the relevant sectors and types of ecological systems. A view commonly expressed by our informants was the need for more and better monitoring data, both in marine and terrestrial environments. Indeed, **the need for an assessment system that is applicable to regional and local levels was emphasised by most of our informants**. The availability of relevant ecological data with high enough resolution to be used at low geographical levels is found to be a major obstacle in this respect. This would most likely imply raising the budget and efforts for data gathering.

The evaluation panel concludes that several problems arise when stakeholders want to make use of the results of the assessments to inform management and develop actions, but that the PAEC method (with its since long established networks between monitoring-assessment-policy, including direct links to marine planning) is better designed than IBECA for such purposes. We also found great expectations among stakeholders that the assessments would become useful also at sub-national levels of governance.

6. Comparing the Overall Strengths and Weaknesses of the PAEC and IBECA Methods

6. Comparing the Overall Strengths and Weaknesses of the PAEC and IBECA Methods

This chapter addresses the evaluation mandate by comparing the two assessment methods. It ends with two summarising tables that compare the scientific and technical properties as well as the applicability of the two methods as basis for decision-making in policy and management.

Overall aim with the assessment

Both the PAEC and IBECA methods are designed to reach a qualitative conclusion on the ecological condition based on the same concept of 'intact ecosystem' as the reference condition. Both base their methodologies on the same seven ecological characteristics of ecosystem condition, put forward by Nybo & Evju (2017): primary productivity, biomass distribution among trophic levels, diversity of functional groups, functionally important species and biophysical structures, landscape ecological patterns, biological diversity, abiotic factors. Indicators are at the core of both methods, linked to the seven ecosystem characteristics.

Both systems are comparable, but they differ in their respective approach to the assessment process. The IBECA method assesses if there is a deviation in data-driven indicators to reference values set by expert judgement. PAEC uses a panel of experts to formulate hypothesis-driven expected deviations of the current ecological condition compared to the reference condition and evaluates the evidence for such deviation.

Definition of deviation from "good ecological condition”

Both PAEC and IBECA define ‘Good ecological condition’ in Norwegian ecosystems by the fact that the structure, function and productivity of ecosystems do not differ significantly from the reference condition, defined as intact ecosystems (Nybø & Evju, 2017, p. 32). Here ‘intact nature’ is where human pressures should be on a scale and of an extent that does not exceed the effect of natural drivers of change.

In IBECA ecological condition is assessed based on a set of quantitative indicators from monitoring data which compare the current ecological condition to a reference condition of “intact nature”. Each indicator is standardised to a scale from 0 to 1, using a piece-wise linear scaling defined by the minimum and maximum values for the indicator and the threshold for a good condition on the original scale, where the latter corresponds to a value of 0.6 on the scale. The standardisations of indicators to unit scales are justified by the group of experts involved in identifying and selecting indicators and are to be applied consistently across systems and scales. The deviation from “good ecological condition” is derived by comparing to 0.6 on individual as well as aggregated scaled indicators. A conclusion from chapter 4 is that the value of scaled indicators and aggregated scaled indicators are difficult to interpret and compare within and across ecological systems.

The PAEC defines a “good ecological condition” as the condition of “an intact ecosystem”, which is a system largely unimpacted by modern industrial anthropogenic activities. More specifically, this can be the state of the ecosystem at a pre-defined time period, or according to specific criteria such as the absence of local and global human influences (“a pristine state”), or the maintenance of important functional or structural components.

In PAEC a selected panel of experts identify indicators, formulate hypotheses on their directions of change and list their primary drivers (scoping). Next, available data for the indicators are analysed together with an evaluation of the scientific quality of each indicator. The panel then assesses the overall ecosystem condition considering the results from the analysis phase, and if there is sufficient data to assess deviation. After a mandatory peer-review, the conclusion from a PAEC can be (1) no evidence for deviation, (2) evidence for limited deviation, (3) evidence for substantial deviation, or (4) insufficient data to assess deviation from the reference condition. The class “no evidence for deviation” corresponds to a “good ecological condition”.

IBECA defines deviation from “good ecological condition” by specifying thresholds on quantitative indicators, which means that the scientific quality of the definition is related to the quality of data and the scientific base and relevance of the threshold. The meaning of an aggregated indicator is limited to how it was derived, rather than representing a deviation from “good ecological condition”.

In PAEC the experts judge if the current state (according to current data) deviates from the reference state. This evaluation is qualitative and allows experts to consider uncertainties external to quantitative information. A qualitative evaluation can be applied on situations where one cannot identify and estimate a threshold value, but rather focuses on the type and direction of changes away from the reference condition that can lead to changes with significant effects on the ecosystem. A quantitative evaluation can be used as support, e.g. when the experts decide to identify and estimate a threshold value for a change away from the reference condition.

Due to these differences in methodology, indicators with poor or no quantitative data on the reference condition can be handled by PAEC (although the validity will be low), whereas IBECA cannot because the reference condition requires quantitative evidence.

Compared to IBECA, the PAEC has a more flexible and more meaningful approach to define a deviation from good ecological condition.

The use of existing monitoring data of an assessment method is a strength, considering monitoring data have a high level of evidence, but also a weakness when a method is limited to this type of data, since in practice the availability or resolution of monitoring data is low.

Communication of conclusion in the assessment

The overall assessment of ecological condition results in the IBECA system as a weighted average of scaled indicator values, aggregated per ecosystem characteristic and the overall ecosystem. The scaled indices for the seven characteristics are visualised as numerals in a radar graph, or forest plot (Figure 1), also showing the aggregated values over the seven characteristics or the individual indicators, respectively. In the national assessments conducted so far, each of the seven ecosystem characteristics together with the ecosystem overall are assessed as Good (certainly good), Degraded (probably deviates from good), Very degraded (certainly deviates from good) or Uncertain (Example in Table 1). This classification comes from expert judgement based on the indicators' coverage of the ecosystem characteristics and the level of indicators (for unscaled indicators it is compared to the value for the reference condition) and for scaled indicators it is compared to 0.6). If available, the conclusion is reported together with information about trends for the indicators' unscaled values, and the effects of main drivers on the scaled values of indicators assigned to each characteristic, as well as the indicator for the overall ecological condition (for examples see Framstad et.al. 2022, Tables 3.9 and 4.9).

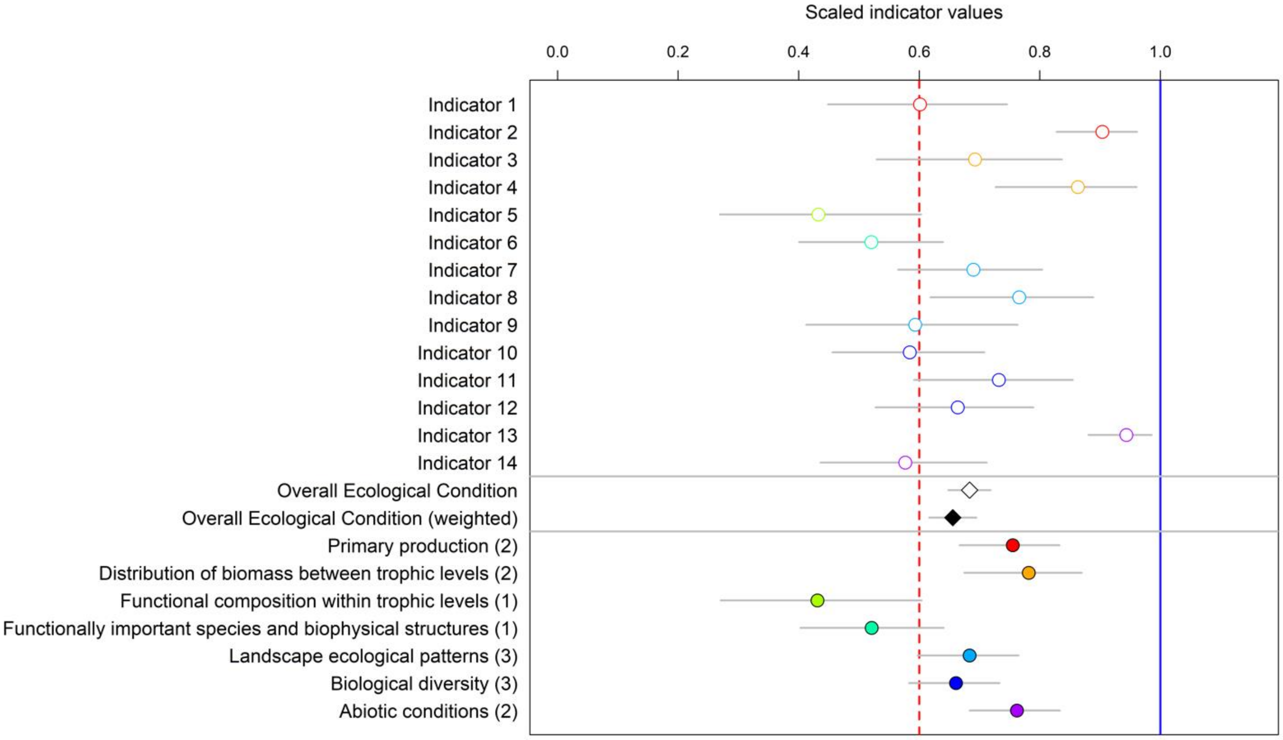

Figure 1. Forest graph of individual indicators, aggregated indicators for the ecological system and per characteristic. Colours are added by the external evaluation panel to show how individual indicators are grouped under the seven ecological characteristics (numbers per group in brackets). Values are fictive and meant for illustration.

Table 1. Overall assessment of the reliability of the results in Figure 1 and based on indicators coverage of the ecosystem characteristics, level and trend in indicator values. Values and judgements on reliability, as well as the conclusions on the ecological conditions, are fictive and meant for illustration.

Ecosystem characteristics	**Condition value**	**Indicators**	**Indicator values - Levels**	**Indicator values - Levels**	**Condition**
Primary production	0.76	Insufficient	Small deviation	Positive	Good
Distribution of biomass between trophic levels	0.78	Insufficient	Some deviation	Negative	Good
Functional composition within trophic levels	0.43	Insufficient	Large deviation	Positive	Very degraded
Functionally important species and biophysical structures	0.52	Insufficient	Small deviation	Negative	Degraded
Landscape ecological patterns	0.68	Insufficient	Small deviation	Negative	Uncertain
Biological diversity	0.66	Insufficient	Small deviation	Negative	Good
Abiotic conditions	0.76	Insufficient	Small deviation	Positive	Uncertain
Overall Ecological Condition	0.68	Insufficient	Small deviation	Negative	Good

IBECA communicates the numerical values of the individual indices, aggregated indices on the seven characteristics and the aggregated overall condition, which could be understood as how close the system is to a good ecological condition. These numbers are part of the bases to make the assessment of the condition, but how to go from the numbers including quantified uncertainty ranges to the conclusion on the overall ecological condition is not made clear. The judgements on the reliability of the basis for the values of indices (presented in the forest plot Figure 1) are presented together with the numbers of the aggregated indicators in a separate table and the assessment of overall ecological condition (fictive example in Table 1). The IBECA system focuses on visualising the values of scaled and aggregated indicators, where qualitative conclusion and summary of uncertainty is presented in text or tables.

In PAEC, the scientific panel concludes regarding the condition of the ecosystem, based on a judgement of ecological condition for each characteristic expressed as Evidence for Substantial deviation, Limited deviation, No deviation, or Insufficient evidence together with a rating of the coverage of data as adequate, partially adequate and inadequate. The conclusion on each of the ecological characteristics is based on judgements on the evidence for phenomenon (Insufficient, None, Low, Intermediate, High), validity of phenomena (Low, Intermediate, High), the data coverage and regional consensus for each selected indicator (Figure 2a). The aggregation of the ecological condition per ecosystem characteristic is summarised in a hierarchical circular graph, which also shows the coverage of phenomena/indicators behind each characteristic (Figure 2b). From this, the overall assessment of ecosystem condition is justified.

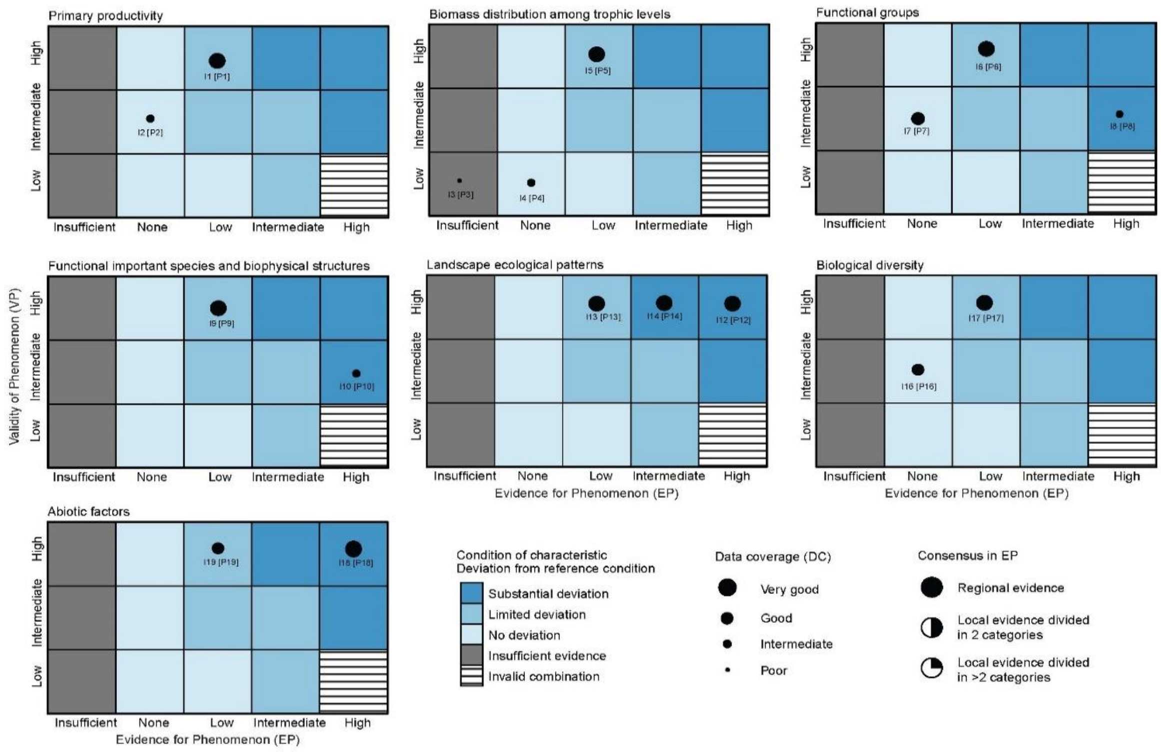

Figure 2 a) Overview of assessments of deviation from reference condition on a categorical scale per indicator ("I") and associated phenomenon ("P") under each of the ecological characteristics, based on judgements on evidence for phenomenon and validity of phenomenon, data coverage and consensus between multiple instances of local evidence per indicator. The judgements are fictive and meant for illustration. Figure created by Jane Jepsen.

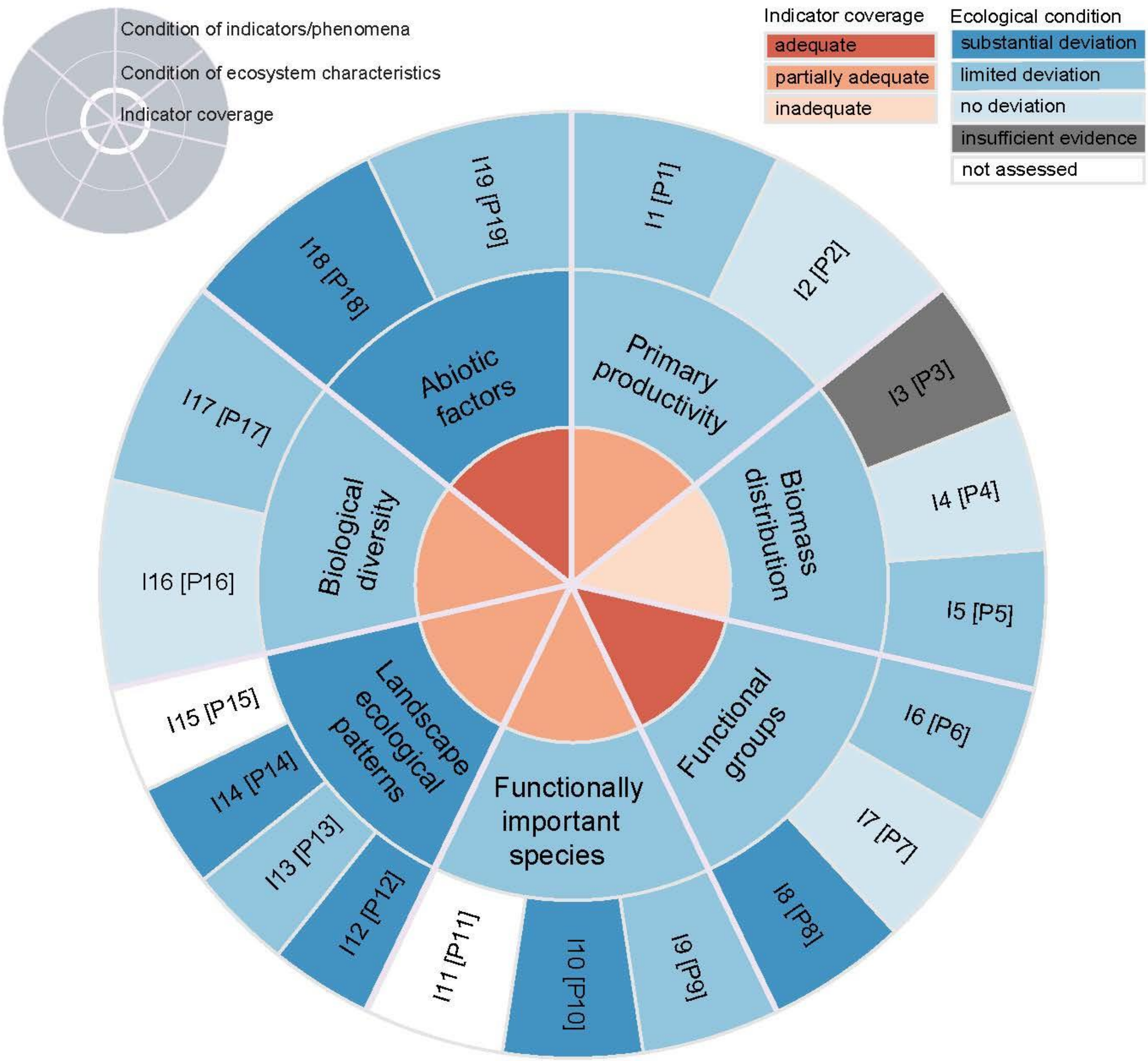

Figure 2 b) A hierarchical circular graph with the assessment of ecological condition of individual indicators ("I") and associated phenomena ("P") and the seven ecosystem characteristics, together with an evaluation of the strength of the scientific basis for the assessment of each ecosystem characteristic (indicator coverage). The judgements are fictive and meant for illustration. Figure created by Jane Jepsen.

The visualisation of the classifications of ecological condition including coverage in the circular hierarchical graph from the PAEC system (Figure 2b) provides a balanced overview of the assessment of phenomena/indicators and the ecological characteristics. The PAEC system communicates the overall assessment of the ecological condition in text using a short ingress which highlights that this is the conclusion reached by the scientific panel.

In addition to presenting an overview the conclusion, both systems include visualisation and tables of indicator values on the original scale in their reports.

We find that both methods visualise the conclusion of the assessments through figures that aid the reader's understanding.

Selection of indicators

According to the mandate given to Nybo & Evju in (2017), the technical system should initially be established for ecosystems at county/region level, or other professionally based, appropriate levels. The assessments made so far have been at both regional and national level. In principle, IBECA and PAEC can be implemented at finer resolutions, but an impediment is data scarcity. If well covered, indicators used by IBECA or PAEC can be derived at finer scales and aggregated to coarser scales. For both systems, new qualitative judgements are needed based on aggregated variable values at the coarser level (see e.g. Framstad et al. 2023). The above concluded flexibility with the PAEC in the problem formulation and transparent use of expert judgement, makes it easier to develop a PAEC assessment for any suitable spatial and temporal scale without being limited to data coverage. Considering resources required for structured approaches of expert judgement, there is a trade-off between the increased number of judgements needed for undertaking assessments at lower scales and the benefit for assessing ecological condition and management.

There is no description in the technical documents on IBECA presented to the evaluation panel (published in English) on the procedure for how indicators are identified and selected. Requirements for their inclusion are that they must be responsive to five identified external pressures, reference values must be possible to quantify as well as limit values and min/max values. Requirements for data in the IBECA method include that they must be spatially representative for the ecosystem in question, reliable with respect to quality, and retrieved from professional performed data collections in monitoring and research. Data that do not fulfil the requirements can be used as supplementary variables. Ideally, time-series data should be used.

In PAEC, it is stated that indicators should be sensitive to changes, applicable over a large area, valid over a wide range of stress and cost-effective. Indicators may range from state variables that directly denote ecological functions and structures to surrogate indices that have validated indirect relations to such functions and structures. In the reporting, motives must be given for the choice of each indicator. Any general choices made on which data sources to include or exclude should be reported, as well as a complete overview of all individual datasets used.

Compared to IBECA, the PAEC is more transparent about the choice of indicators and critical evaluation of data.

Linking indicators to drivers for change

Both assessment methods link indicators to drivers for change, but the impact of such drivers are not explicitly statistically tested in either of the methods.

In IBECA, causes for change in ecological condition are considered by choosing indicators that describe drivers or responses to drivers as in the DPSIR framework. The IBECA method generates indices for different drivers by aggregating values for indicators related to habitat loss, climate change, pollution/eutrophication /acidification), exploitation/harvesting and invasive alien species, separately. An aggregated indicator for a driver is seen to provide information about what causes a possible deviation from good ecological condition. A limitation is that the scaled indicators describe a current condition and do not explicitly consider the influence of drivers. Methods for inference of trends or changes would be more appropriate to conclude on the influence of a driver of change, but this is often limited by availability of data. There are in applications of IBECA also summarises of what is known about the influence of drivers of change based on additional analysis and available data behind the indicators (see example in Table 1).

A key component in PAEC is the formulation of hypotheses ("phenomena", quantitative or qualitative) describing expected directional change away from the reference condition for each of the indicators as a result of relevant drivers in the system. The scientific basis for each phenomenon is evaluated, reflecting how well the links between drivers and indicators are understood, as well as how well the indicator reflects the condition of the ecological characteristics and the wider ecological significance of the phenomena. The PAEC system does not explicitly relate to the DPSIR framework.

The term "phenomena" used in PAEC is partly difficult to understand and often changes in indicators are more frequently discussed in their communication.

PAEC can include drivers for change in the specification of phenomena, which when addressed by a mixture of quantitative and qualitative methods allows for a stronger scientific basis to evaluate change from drivers compared to comparing aggregating scaled indicators for different drivers as done in IBECA.

Methodology

Modelling and uncertainty

Both IBECA and PAEC rely on quantitative and statistical methods for inference on quantitative data from various sources, but with different requirements on how to account for uncertainty.

IBECA puts emphasis on the direct characterisation of uncertainty from the statistical method and variability in data using quantitative methods. When aggregating, the IBECA method characterises the combined impact of quantified uncertainty associated with each indicator on the aggregated indicator. In IBECA, there is a discussion of the reliability of results, which also can include a qualitative overall assessment of reliability (see Framstad et.al, 2022, Tables 3.9 and 4.9), but weakness due to lower reliability is not considered in the aggregated indicator for the overall ecological condition.

The PAEC method uses a structured "weight of evidence" type of approach to make the overall assessment, considering a qualitative scoring of the individual indicator analyses. This means that in addition to direct uncertainty in data or model-based estimates, the experts are explicitly asked to critically evaluate the scientific strength of each indicator analysis and the extent to which their influence on the phenomena are understood (a type of indirect uncertainty). In PAEC this critical appraisal is summarised in categories based on spatial representativity, temporal representativity, data coverage and indicator coverage (Figure 2a).

Both direct and indirect uncertainty are important to consider in scientific assessment, and PAEC has, compared with IBECA, a more structured and elaborated approach to consider and communicate limitations in the scientific bases (indirect uncertainty) on the conclusion.

In both methods severe data gaps result in lack of indicators or poor scientific basis for an ecological characteristic. In IBECA this results in few indicators to be included in aggregated indicators thereby reducing the validity of the overall indicator. In IBECA, data gaps are reflected in the discussion of reliability of the overall indicator, but it is not clear whether and how lower reliability is considered when assessing the overall ecological condition or conditions pertaining to the seven characteristics. Consequently, it is left to policy makers to judge if the assessment has enough scientific basis and what conclusion to make. In PAEC, the experts must decide if data gaps should result in the classification “insufficient data to assess deviation from the reference condition” or “no evidence for deviation” (i.e. a good ecological condition). Here it is important that the meanings of these outcomes are communicated carefully, since concluding “no evidence for deviation” can be misunderstood as no change when there is not enough evidence available for the assessment.

Selection and use of expert judgement

Judgements made by experts are used in both assessments, but in a more or less formal way.

IBECA uses experts to define reference values and limit values for good ecological condition, to calibrate the scaling of the indicators, to support the characterisation of uncertainty in indicator values, and to weight or propose the method to weight the indicators when aggregating. The PAEC method uses panel experts for all steps of the assessment, following a structured process, finalised into a group judgement. This judgement is the opinion of the panel. When a panel member declares a reservation, this is reported, and he/she is named.

In IBECA, the assessment is performed by a rather small group of experts at a few institutes, led by NINA. Criteria for selection of institute(s) or experts are not given. In PAEC, the work is led by HI for the marine environment and by the NPI for the Arctic tundra while including institutes with scientific members from different institutes and universities. Criteria for panel members in PAEC are that they should have 1) good enough knowledge of the focal ecosystem to formulate and validate phenomena, identify indicators and data sources, and/or have 2) expertise on quantitative methodology (study design and statistical modelling). No criteria for selection of institute(s) to lead the work have been presented.

In this report we identify that expertise in social science to understand societal and human values when setting thresholds for indicators and expertise in structured expert judgements are needed, but this is currently not asked for in the description of the methods.

A strength of PAEC is that it embeds expert judgement as part of the assessment, but both methods could be more transparent and formal in their selection and use of expert judgement.

Reporting

Reporting in applications of both assessment methods has so far been extensive, with multi-page publications including much technical detail, and with some reports only published in Norwegian. Data and codes are not available to a satisfying extent, but IBECA has documentation with shareable code on public repositories. Both systems move towards adopting principles for open and reproducible science, which is a requirement to meet demands for transparency.

PAEC reports must be peer-reviewed by other experts or stakeholders, before publication. Applications of both assessment methods have been published in peer-reviewed scientific journals: IBECA (Jakobsson et al. 2020, 2021) and PAEC (Jepsen et.al. 2024). A built-in procedure for independent peer review is an asset of a system for scientific assessment.

Both systems face challenges in reporting, not only with respect to the amount of information but also to meet demands for reproducibility and transparency. The evaluation panel went through the reports several times to try to grasp and interpret them correctly, which means that important points can easily get lost in the details. Still, we found that **PAEC has stronger characteristics with respect to peer-review, thereby allowing for critical evaluation of independent experts on the work. We also found that PAEC assessment reports are somewhat more accessible for a broader audience compared with those conducted with the IBECA method. But both methods would need to be summarised and communicated in ways that are easier to comprehend for laymen and relevant stakeholders.**

Legitimacy and usefulness for policy making

Repeatability and flexibility over time

Both methods can be repeated and consider newly available data, with a suggested cycle of every fifth year, which coincides with the cycle of Nature Index data updating. Both methods can adapt to new scientific knowledge concerning the choice of indicators, their drivers, reference/threshold values, and (if applicable) the principle for scaling.

A repetition of an PAEC assessment can be done based on a new expert panel, however with a revision of the scoping as well as any consideration of new data. The PAEC method includes the step of selecting indicators/phenomena in every assessment, and therefore including new data can be done. The first time an assessment is done will be the most demanding, while repeating an assessment is expected to be easier every new time. The efficiency of repeating a PAEC depends on the choice of experts, what amount of work that can be done internally and what must be supported by independent experts.

PAEC suggests viewing a repetition as a re-evaluation of the previous assessment, where there is also made a summary of the differences compared to the previous assessment. We find that an added value of the PAEC system is that they have a strategy to report the difference in the repeated assessment compared to previous assessment. This will help decision-makers understand why the classification on ecological condition might have changed.

The ambition of IBECA is that repeated assessments should be possible also for new expert groups. The underlying assumptions in indicator selection and setting of values for upper/lower reference conditions and for limits of good condition, as well as in qualitative assessments of the results will then have to be properly documented in the published scripts for assessments of each ecosystem. We find that a repetition of IBECA for new experts is currently difficult since embedded assumptions and judgements are not fully explained or documented in the IBECA reports reviewed in this mandate. Considering repeating an assessment, the IBECA recommends to also consider a possible revision of indicators, reference values or values for good ecological condition (for example to consider new data), but does not give any recommendation if the new assessment should be compared to the old one.

A well calibrated IBECA assessment can - when applied on data from the same sources and of the same quality - provide consistent and comparable assessment in time and space, and thereby be updated with efficiency.

Both systems can be repeated, but PAEC has a stronger strategy to communicate and justify differences from one assessment to another compared with IBECA.

Compliance with established methods to assess ecological condition

IBECA has similarities to the EU Water Framework Directive in defining reference values based on “intact nature”, scaling of indicators, and using > 0.6 as threshold for good ecological condition. The method is also part of the development of UN and EU ecosystem accounting systems SEEA EA. PAEC has many similarities to the approaches for scientific assessments conducted as part of IPCC and IPBES as well as the terrestrial COAT programme in the Arctic Tundra.

Both systems are following established but different assessment frameworks.

Link to management plans

Although the two systems do not have an explicit requirement to be used to support management, we have evaluated to what extent they can be linked to management plans. One aspect is that the assessment system can evaluate the impact of management actions or identify what management actions to propose. Their usefulness depends on what type of drivers or management variables that are considered within the assessments, or if the outcome of the assessment can be used to prioritise conservation efforts, e.g. via protected areas. These things are not covered in any of the assessment methods in a systematic and enough detailed way. The links to drivers of change are looking into the past rather than into the future, although likely future developments of drivers and their potential impacts on indicators are discussed in reports of both methods. Taken together, it is difficult to use the results of the assessments to develop proposals for policy or management measures to improve the ecological condition.

Summary comparing the scientific strengths and weaknesses

We display the strengths and weaknesses of the two methods in the following tables. The first summarises how we evaluate the scientific and technical properties and the second shows how we evaluate the methods in relation to their applicability as basis for decision-making in policy and management. We add a comment to explain what we consider as an overall judgement for each of the aspects that relate to our initial proposal for a well-functioning assessment system for ecological condition (see Chapter 2 under ‘Positioning this evaluation in relation to our mandate’).

Tab 2 page 1	Main assessment approach	Statement of ecological condition	Indicators	Thresholds	Reference condition	Data	Use of expert judgements
PAEC	An approach based on hypothesis formulation, comparing reference conditions to current temporal trends using data analysis and modelling in combination with expert judgement, and with final assessment made by the expert panel.	Stated as categories: Substantial deviation, Limited deviation, No deviation from reference condition, or Insufficient evidence	Selected based on formulated hypotheses (phenomena) regarding drivers of ecosystem change	May be considered for indicators where time-series data include the reference state (e.g. climatic data)	Stated in a normative/qualitative way based on different and explicit sources of scientific evidence. When data include observations from the reference state (e.g. climatic data), reference condition can be expressed quantitatively.	Time-series data based on monitoring, for which temporal trends (and in some cases also deviation from reference condition) are analysed.	Used in almost all steps of the assessment process, following a structured process, finalised into a group judgement.
IBECA	An indicator-based approach using quantitative comparisons between reference values and current states (snapshots). The final, qualitative classification is based on numerical values.	Stated as categories: Good, Degraded, Very degraded or Uncertain. The threshold for good ecological condition is set to 0.6, on a scale 0-1.	Selected based on knowledge of structures and processes that determine the state of the ecosystem	A foundation in the method and used for all indicators. A threshold for good ecological condition is set to 0.6 on a scale 0-1.	Always expressed quantitatively, based on various sources of scientific evidence and expert judgement, however not fully explicit how the values are derived	Snapshot data for reference state and current state mostly based on monitoring. Time-series data can be used as support	Used in several steps such as in the selection of indicators, in the determination of some reference values and limit values, but with no formal process for working with experts.
Comment	Although using similar type of ecosystem data (mostly monitoring data), there are fundamental differences in the methods regarding design, analyses and reporting	Both methods result in a statement regarding if an ecosystem has good ecological condition or not. PAEC has higher resolution in statements. The usefulness of the respective types of statements depends on the users' requirements	Motives for selection of indicators are more clearly expressed in PAEC than IBECA.	A drawback with IBECA is that a scientifically based motivation for 0.6 as a threshold is lacking.	A drawback with IBECA is that uncertainty in reference values is not considered in uncertainty of the scaled or aggregated indicators.	The data used by PAEC and IBECA are of similar type, usually based on monitoring. Limited data coverage may be an impediment for both methods	Expert judgement is used in PAEC in a generally transparent and clear way. When used in IBECA, transparency and clarity is lower, and occasionally not reported at all.

Table 2. (split into 2 pages) Comparing the PAEC and IBECA methods in relation to scientific and technical criteria for an efficient assessment system

Tab. 2 page 2	Links to human impacts	Scalability	Assessment of uncertainty	Transparency in methodology	Reporting	Repetition over time	Compliance with established methods
PAEC	In the hypotheses (phenomena), human impacts are a main component, discussed for each indicator	Can be implemented at finer resolutions, provided sufficient data coverage. New qualitative judgements are needed based on aggregated variable values at coarser level	Uncertainties in estimates are characterised using quantitative methods. The scientific strength of each phenomena is judged qualitatively in a structured way and considered when assessing the phenomena.	PAEC follows a structured procedure with specific requirements on methodology and reporting at each step.	The PAEC method has been published scientifically. Reports are extensive and technical. PAEC reports must be peer-reviewed by other experts or stakeholders	Repetition of PAEC assessments possible with new expert judgement and might require formulation of new hypotheses. Recommends showing differences between re-evaluations.	PAEC has many similarities to the approaches for scientific assessments conducted as part of IPCC and IPBES. PAEC experts are engaged in OSPAR and ICES assessments.
IBECA	IBECA uses the DPSIR framework of human impacts for each indicator and aggregates to ecosystem level	Can be implemented at finer resolutions, provided sufficient data coverage. New qualitative judgements are needed based on aggregated variable values at coarser level	Uncertainty in scaled indicators is characterised from the statistical method or variability in data using quantitative methods. Uncertainty in aggregated indicators is characterised by propagating uncertainty from scaled indicators when aggregating. The reliability of indicators is discussed in a general way.	IBECA follows a structured procedure with specific requirements on methodology at each step, but expressed with less clarity in the reports	The IBECA method has been published scientifically. Reports are extensive and technical.	For assessments not in need of updating of indicators, revisions only need inclusions of new monitoring data. If a revision includes change in indicators, new descriptive and analytical work is required.	IBECA resembles the EU Water Framework Directive in defining reference values based on “intact nature”, scaling of indicators, and using >0.6 as threshold for good ecological condition. The IBECA method is compatible with UN and EU ecosystem accounting systems SEEA EA
Comment	For both methods natural variation (variability) in indicators may be hard to separate from change caused by human impact.	Judgement-based conclusions restrict aggregation of PAEC assessment from different areas more than IBECA. For both methods, lack of data with high coverage of ecosystem extent and high resolution prevents application at small scale.	IBECA and PAEC both rely on quantitative and statistical methods for inference on quantitative data but PAEC has a strength in also making a formalised, qualitative judgements of the scientific strengths of indicators.	PAEC has a higher transparency in how the assessments are conducted.	IBECA reports have somewhat less clarity than PAEC reports. There is a need for a strategy for popular communication to stakeholders	A drawback with PAEC is that revisions will likely include different panel members, potentially affecting the design of the assessment. IBECA is easily updated when using the same selection and thresholds for scaling for indicators.	Both methods are in line with methodology used and accepted internationally but in different ways and contexts.

	Clear aim and well communicated	Allowing stakeholder input	Which stakeholders are relevant including ITK	Open discussions about improvement	Easy to understand judgements for laymen	Legitimacy among relevant stakeholders
PAEC	Clear aim for marine ecosystems: feeds directly into marine management planning processes. Well communicated through established stakeholder groups. Not analysed for Arctic tundra ecosystems.	In principle yes, but in practice not (so far)	Not clear on which stakeholders are relevant to include	Not as far as we have seen	Reasonably well justified and transparent results	High
IBECA	Different conceptions on aim among stakeholders; Reports difficult to grasp, but good with outreach via webinar(s)	No	Not relevant as method is focusing on the quantitative part	No	Difficult to follow some steps in the process of establishing 'good ecological condition'	Low for economic forestry; high for environmental organisations; unknown for mountains as we did not explicitly examine stakeholder views on this ecosystem.
Comment	PAEC results connect to planning processes while IBECA does not although this is expected by some stakeholders	Both methods could improve	Both methods could improve; Indigenous and local knowledge input not mentioned	Unclear to us whether PAEC does this via the Monitoring Group	PAEC much easier to understand how they get to the result compared with IBECA	PAEC applied to marine ecosystems benefits from high trust due to well-established connections between science and policy through the marine management planning processes

Table 3: Comparing the PAEC and IBECA methods in relation to their applicability as basis for decision-making in policy and management

7. Conclusions and Recommendations

7. Conclusions and Recommendations

In this chapter, we summarise our findings according to the mandate for this external evaluation, ending with our recommendations. We begin by making some general conclusions and observations.

Overarching conclusion

The essence of our mandate was to compare the PAEC and IBECA methods, which have been developed for the same purpose but in partly different institutional and policy contexts. We find that neither of the two methods follow all our proposed scientific and management criteria for an efficient assessment system for ecological condition, but that the PAEC method is better designed overall to this aim compared with the IBECA method. Since our knowledge of the processes influencing the functioning of the different ecosystems are incomplete, and indicators must be selected based on data availability and best available knowledge, we fully acknowledge the need for structured and well-documented expert judgements. In this respect, we find that the PAEC method outweighs IBECA. IBECA relies heavily on quantitative indices and aggregation of data with comparatively low transparency and high uncertainty in relation to the setting of a normative threshold of 'good ecological condition'. Inevitably, to evaluate changes over time and critical thresholds not to be exceeded will require insightful expert judgement from different perspectives and needs to be presented in a transparent manner.

General findings and observations

- There are expectations from Norwegian national authorities that the assessment system should be used for several purposes, such as designing policy goals and formulating management objectives but also to connect to international processes for biodiversity reporting and protection. From our interviews with stakeholders, we learned that there are widely different views on the purpose of the system, and thus we consider it fundamental that the aim is unambiguously expressed. It will also be important to firmly establish that the assessments as such should be separated from goal-setting processes.

- A fundamental assumption for the assessment system is that it should be able to clarify what is meant with a 'good ecological condition'. However, we find that the combined effect of using an "intact" ecosystem as a goal, together with the partly non-transparent way that the individual reference values have been set, makes it difficult to know what a good ecological condition is compared with. The evaluation panel wants to bring attention to the embedded social and political values in defining such a condition as well as linking it to a normative value of 'good ecological condition'. A possibility to remedy this problem would be to include not only natural science experts in the assessments, but also social science experts who can draw on socio-economic and cultural conceptions of these concepts and link the establishment of a 'reference condition' to a broader discussion of what status is desired for the different types of ecosystems.

- We also find that lack of indicators is the case for all ecosystems, and that a significant effort would be required to rectify this with improved ecological data gathering. Even if certain processes are sufficiently known to select an appropriate indicator, that will rarely be the case for all those involved and it will be very difficult to assess the state of an ecosystem by combining all the indicators. In other words, the IBECA approach will require the insights of experts as is done in PAEC.

- We further note that soil organisms of terrestrial ecosystems need more attention in the future due to their critical role in above-ground ecological processes. Measurements of soil organisms might become increasingly critical in coming years.

- Albeit that the original mandate from KLD (Appendix in the 2017 report) states that the county/regional level should be addressed, with time, there seems to be a successive orientation towards the national level, although the regional/county level is targeted for the three sea regions, and for forests and mountains. Thus, there is a need for clarification regarding relevant and possible levels to address. Avoiding assessment at lower administrative levels limits the possibility to analyse drivers for change and management responses in different regions, and thus restricts the method's applicability to proposing scientifically based restoration measures for specific landscape areas.

- We find widespread agreement across stakeholders that action plans are needed as follow up of the assessments. Many of our interviewees (both stakeholders and assessment experts) express expectations that the assessment system should be useful not only at national level, but also for conducting regional and local assessments to feed into e.g. land-use planning and natural capital accounting at municipal level. This would require adapting the assessment method to such needs by e.g. improved data gathering as well as modelling that is less coarse compared to the present situation. We find that while technical improvements in data gathering can most likely be achieved with remote sensing and AI techniques, quality control in the field is often necessary to circumvent poor interpretation and relevance of those data. Strengthened – and higher resolution – data gathering will therefore be key to useful application of assessment methods below the national level.

- Of the five assessments conducted so far, four are made using the PAEC method (Norwegian Barents Sea Shelf Ecosystem, North Sea Shelf Ecosystem, Norwegian Sea Pelagic Ecosystem, Norwegian Arctic tundra) and two using the IBECA method (forest and mountain ecosystems). The imbalance in number of assessments complicates comparisons between methods and consequently we have not emphasised this part of the mandate in our evaluation. Still, we have put weight on how these assessments have been perceived and used as input to policy and decision-making and have compared the two methods in this respect.

- None of the PAEC and IBECA assessment processes so far have involved stakeholders since this was explicitly mandated by the MD, although the PAEC method protocol allows for such interaction. At the same time, especially the IBECA forest assessment has revealed low legitimacy among forest sector stakeholders for the assessment method. We find that engagement with relevant stakeholders at appropriate stages of the assessment process could help improve the relevance and legitimacy of the assessments.

- We note that neither in the relevant policy documents nor in any of the evaluated assessment reports is there any mentioning of the importance of including Indigenous and local knowledge in assessments of ecological condition or in the subsequent policy processes. We find this to be surprising given the emphasis in CBD and IBPES to take such knowledge into account in the biodiversity governance systems, of which monitoring is part.

- We perceive our task to include the aspects of cost-efficiency and the time interval for updating, since this was mentioned in the original 2017 mandate. Still, for both those aspects, the evaluation panel abstains from making any judgements due to insufficient background information. Regarding cost-efficiency we obtained some information on time and resources spent for the pilot studies and the first assessments but without enough detail to allow analysis. For the updating, 5 years is interval for the Nature Index, but it was mentioned by IBECA experts that 3 years is planned for some indicators within the EU ecological accounting system. Hence, we leave this to the Norwegian experts and ministries to reflect upon.

- All stakeholders agreed that the presentation of results is key, but also that the reports so far have been too thick and technically detailed for many to be able to read. The evaluation panel concludes that the current communication outlets for the assessments carried out so far are not fit for reaching out to the relevant stakeholders.

- The evaluation panel finds that some of the problems that we have detected with the current assessment system can be modified by the expert groups in charge of the assessments, but that other issues are more complicated and will require specific action by the Ministry.

Finally, we present our recommendations both to the PAEC and IBECA methods and to the Ministry for Climate and Environment for the future direction of the assessment system for ecological condition.

Recommendations for both assessment methods

- Ensure that expert judgement is done according to best practices. Allocate work to internal staff to reduce the workload on external independent experts. Train experts and staff.
- Encourage IBECA and PAEC to express ecological condition in a common “currency” to facilitate wider communication with policy makers
- Further investigate and explain the connection between relevant natural and anthropogenic drivers for change and selected indicators to facilitate the development of management options
- Develop and implement the peer-review process
- Improve stakeholder involvement, e.g. through consultation
- Find a common format for user-friendly summaries of results

To improve the PAEC system

- Explain better the judgements of ecological condition made in relation to reference states and time-series
- Explain better how stakeholders can be involved
- Use processes for expert judgement that reduces biases
- Clarify the distinction between ‘No evidence for deviation’ and ‘Evidence for no deviation’

To improve the IBECA system

- Give justifications for expert judgements, provide more transparency, and use processes for expert judgement that reduces biases
- Engage stakeholders in the process
- Write reports in a more structured way
- Clarify which sources of uncertainty that have been considered in the derivation of uncertainty intervals for the indicators, and what impact any additional sources of uncertainty might have on the indicator value and how these are considered in the conclusion
- Conclude with statements on the condition with uncertainty

Recommendations to the Ministry of Climate and Environment

- Be very clear on the aim with the assessments (such as management plans, nature accounting, international reporting) and that the system generates assessment rather than goalsetting
- Clarify which possible governance levels (national, regional, local) to address: increase data gathering and monitoring efforts of terrestrial as well as marine ecosystems, ensuring long time series data
- Investigate further how a system could be developed for local application (fylke, municipalities) for various purposes
- Include also Indigenous and Local Knowledge (ILK) and social science/economics expertise in the assessments
- Consider creating one or two terrestrial expert groups (for Arctic tundra, mountains and forests) on monitoring similar to “Overvåkningsgruppen” (Monitoring Group) for marine environments
- Consider creating stakeholder groups that would link to such monitoring groups on terrestrial ecosystems resembling that for the marine ecosystems “Faglig forum for norske havområder” (Management forum for the Norwegian sea areas) to discuss terrestrial management plans
- Require simpler reporting to stakeholders (reports are now very technical and difficult to read) and make a (financed) communication plan for the assessments
- Allocate (new) resources to interdisciplinary research on the scientific and technical improvement of the two assessment systems as well as their relevance for input to current policy and management development at different governance levels

References

Arneberg, P. et al (2023a) Panel-based Assessment of Ecosystem Condition of the Norwegian Sea Pelagic Ecosystem. Report 16, M-2510|2023, Institute for Marine Research.

Arneberg, P. et al (2023b) Panel-based Assessment of Ecosystem Condition of the North Sea Shelf Ecosystem. Report 17, M-2509|2023, Institute for Marine Research.

Baker, S. & K. Eckerberg (2016) Ecological restoration success: a policy analysis understanding. Restoration Ecology, 24(3), 284–290.

Beck, S., Borie, M., Chilvers, J., Esguerra, A., Heubach, K., Hulme, M., . . . Miller, C. (2014). Towards a reflexive turn in the governance of global environmental expertise. The cases of the IPCC and the IPBES. GAIA-Ecological Perspectives for Science and Society, 23(2), 80-87.

Conference of the Parties to the Convention on Biological Diversity (2022) Kunming-Montreal global biodiversity framework, Montreal, Canada.

EEA (1999) Environmental indicators: Typology and overview, Technical Report 25, https://www.eea.europa.eu/publications/TEC25

Framstad, E., Kolstad, A. L., Nybø, S., Töpper, J. & Vandvik, V. (2022). The condition of forest and mountain ecosystems in Norway. Assessment by the IBECA method. NINA Report 2100, Norwegian Institute for Nature Research.

Framstad, E., Czúcz, B, Schartau, A.K., Simensen, T., Nybø, S. & Sandvik, H. (2023). Naturregnskap og økologisk tilstand: Samsvar mellom fagsystemet for økologisk tilstand, vannforskriften, FNs rammeverk og EUs forslag til naturregnskap. NINA Rapport 2327, Norsk institutt for naturforskning. (In Norwegian with English abstract).

Hilding-Rydevik, T., Moen, J. & Green, C. (2018) Baselines and the shifting baseline syndrome: exploring frames of reference in nature conservation, in: Crumley et al (eds) Issues and concepts in historical ecology: the past and future of landscapes and regions. Cambridge University Press, 112-141.

Jakobsson, S., Evju, M., Framstad, E,. Imbert, A., Lyngstad, A., Sickel, H., Sverdrup-Thygeson, A., Töpper, J.P., Vandvik, V., Velle, L.G., Aarestad, P.A. & S. Nybø (2021) Introducing the index-based ecological condition assessment framework (IBECA). Ecological Indicators 124, 107252.

Jakobsson, S., Töpper, J.P., Evju, M., Framstad, E., Lyngstad, A., Pedersen, B., Sickel, H., Sverdrup-Thygesen, A., Vandvik, V., Velle, L.G., Aarestad, P.A. & S. Nybø (2020) Setting reference levels and limits for good ecological condition in terrestrial ecosystems – Insights from a case study based on the IBECA approach. Ecological Indicators, 106492.

Jepsen, J.U., Arneberg, P., Ims, R.A., Siwertsson, A. & Yoccoz, N.G. (2020). Panel-based Assessment of Ecosystem Condition (PAEC) – Technical protocol version 2. NINA Report 1890, Norwegian Institute for Nature Research.

Jepsen, J.U. et al (2022) Panel-based Assessment of Ecosystem Condition – a methodological pilot for four terrestrial ecosystems in Trøndelag. Report 2094, M-2190|2021, Norwegian Institute for Nature Research.

Jepsen, J., Arneberg, P., Ims, R.A., Siwertsson, A., Yoccoz, N.G., Faucauld, P., Pedersen, Å.Ø., van der Meeren, G.I., & C. von Quillfeldt (2024) Panel-based assessment of ecosystem condition as a platform for adaptive and knowledge driven management. Environmental management, 74, 1020–1036 .

Meld. St. 35 (2023–2024) Bærekraftig bruk og bevaring av natur, Norsk handlingsplan for naturmangfold, Melding til Stortinget.

Miljødirektoratet & Landbruksdirektoratet (2023) Kunnskapsgrunnlag om økologisk tilstand i norsk skog og utredning av tiltak –oppdrag til Miljødirektoratet, Landbruksdirektoratet og flere etater, Rapport M2597.

Millenium Ecosystem Assessment, 2005. Ecosystems and human well-being: current state and trends. Edited by Rashid Hassan, Robert Scholes, Neville Ash.

Nybø, S. & Evju, M. red (2017) Fagsystem for fastsetting av god økologisk tilstand. Forslag fra et ekspertråd, Norwegian Institute for Nature Research.

Nybø & Evju, red (2017) A technical system for assessing good ecological condition. Recommendations from an Expert Committee: English translation, Norwegian Institute for Nature Research.

Nybø, S., Framstad, E., Jakobsson, S., Evju, M., Lyngstad, A., Sickel, H., Sverdrup-Thygeson, A., Töpper, J., Vandvik, V., Velle, L.G. & Aarrestad, P.A. (2019) Test av fagsystemet for økologisk tilstand for terrestriske økosystemer i Trøndelag. NINA Rapport 1672. Norsk institutt for naturforskning.

Nybø, S., Framstad, E., Jakobsson, S., Evju, M., Lyngstad, A., Sickel, H., Sverdrup-Thygeson, A., Töpper, J., Vandvik, V., Velle, L.G. & Aarrestad, P.A. 2019. Test of the system for assessing ecological condition for terrestrial ecosystems in Trøndelag. NINA Report 1672. English translation of selected chapters. Norwegian institute for Nature Research.

Nybø, S., Framstad, E., Jakobsson, S., Töpper, J. & Vandvik, V. (2020) Økologisk tilstand og andre verktøy for å vurdere naturkvaliteter i terrestriske miljø. Datakilder og forvaltningsmål. NINA Rapport 1902. Norsk institutt for naturforskning.

O’Hagan, A. (2019) Expert Knowledge Elicitation: Subjective but Scientific. The American Statistician, 73 (Sup 1), 69-81.

Pedersen, Å.Ø. et al (2021a) Norwegian Arctic Tundra: a Panel-based Assessment of Ecosystem Condition, Report 153, Norwegian Polar Institute.

Pedersen Å.Ø. et al (2021b) Panel-based Assessment of Ecosystem Condition (PAEC) as a Knowledge Platform for Ecosystem-based Management of Norwegian Arctic Tundra, Brief Report 056, Norwegian Polar Institute.

Siwertsson, A. et al (2023) Panel-based Assessment of Ecosystem Condition of Norwegian Barents Sea Shelf Ecosystems. Report 14, M-2511|2023, Institute for Marine Research.

Töpper J. & Jakobsson S. (2021) The Index-Based Ecological Condition Assessment (IBECA) - Technical protocol, version 1.0. NINA Report 1967, Norwegian Institute for Nature Research.

Appendices

Appendix A. Mandate for the evaluation of the scientific basis of the Assessment System for Ecological Condition

(Fagsystem for økologisk tilstand)

Background

Following the white paper Meld. St. 14 (2015-2016) *Nature for life – Norway's national biodiversity action plan*, a scientific advisory group appointed by the Ministry of Climate and Environment published in 2017 the report "*A technical system for assessing good ecological condition – recommendations from an expert committee",* which presents a set of ideas and proposals for an assessment system for the ecological condition of Norwegian ecosystems. The Water Framework Directive (WFD) has previously been implemented through the national policy; thus, the new assessment system was aimed at ecosystems not covered by the WFD, i.e. terrestrial and marine ecosystems exclusively.

The report presents a technical system for the determination of good ecological condition and builds upon and supplements already established classification systems, e.g. the Nature Index, Marine Management Plans, Red Lists and Alien species lists, etc.

Two assessment methods are used to assess ecological condition:

- **Panel-Based Assessment of Ecosystem Condition (PAEC)**
  This method employs studies of individual indicators over time using statistical analyses. An expert panel evaluates whether changes represent deviation from good ecological condition.
- **Index-Based Ecological Condition Assessment (IBECA)**
  This method is based on indexes calculated from the current status of indicators compared to an intact ecosystem. The indexes indicate ecosystem condition on the scale from intact ecosystem to totally degraded ecosystem.

Both methods have since been employed for various ecosystems. As of today, assessments of the terrestrial ecosystems: forests, mountains and Arctic tundra as well as marine ecosystems have been conducted. Before these assessments, the system was developed and tested through several pilot studies (published reports). The pilot studies and assessments have been conducted based on existing data, but The Norwegian Environment Agency has since specified that the knowledge base should be improved for better assessments.

Purpose of the evaluation

The Ministry of Climate and Environment has asked for an evaluation of the scientific basis of the Assessment System for Ecological Condition. In the evaluation, scientific strengths and weaknesses of the two assessment methods should be pointed out, as well as of the assessments conducted this far. The suitability of the two assessment methods for evidence-based policymaking is also important. The evaluation should be based on the mandate set for the scientific advisory group responsible for the 2017 report, see Appendix x Mandate for Expert Council on Ecological Status.

Tasks for the evaluation committee

Evaluate the scientific basis of the Assessment System for Ecological Condition - including:

1. Evaluate the scientific *strengths and weaknesses of the two assessment methods.*
2. Evaluate the scientific *strengths and weaknesses of the assessments conducted since 2017.*
3. Present possible *improvements and adjustments to the assessment system*, both regarding methods and scientific design, including opportunities and dilemmas.

4. As far as the available information permits, we invite the committee to discuss the role of the two assessment methods as basis for decisions in public administration and policy including any limitations of this knowledge base.

The Evaluation Committee should include other issues they find relevant and give specific recommendations on improvements where relevant/necessary, for example ecosystem accounting and/or other systems that may be relevant for Norway. The Committee will need 6 – 8 experts, with a wide ranging expertise, between method and statistics, terrestrial biology and ecology, marine biology and ecology, as well as humanities and social sciences.

Deliverables and schedule

The Evaluation Committee is requested to make a report of its findings and recommendations for improvements of the Assessment System for Ecological Condition, according to the Tasks for the evaluation panel. The report should include a summary.

The Committee will start the work in March, and deadline for the final report is 1st November 2024. Present a note to the RCN with preliminary assessments and proposals for improvements before June 20, 2024. The committee chair is expected to present the evaluation and its major findings in connection with the handover to the Ministry of Climate and Environment.

Data and methodology

The evaluation should be based on the 2017 advisory report as well as all assessment reports and assessment method tests made available for the evaluation panel, see an overview in Appendix 1 We plan for a meeting early in the process where the Evaluation Committee meets the 2017 advisory group and scientific authors of assessments and method tests.

The evaluation committee may also include other relevant peer reviewed, scientific articles than those used by the advisory group. The Evaluation Committee may request translation of other relevant documents and further information as needed. The Evaluation Committee is free to contact additional expertise when needed and should notify the RCN if they do so.

Organisation and work process

The Evaluation committee will have several virtual/video meetings and at least two physical meetings, one early in the process and one in the last part of the work. The evaluation Committee may, after its first meeting, suggest clarifications in the mandate.

The Research Council will have the administrative responsibility for the following work:

- Be the main contact point in the process for the committee chair
- Clarifying any ambiguities and/or approve any changes in the mandate
- Help facilitate the process, and plan a secretariat function in cooperation with the committee chair
- Host physical meetings at the Research Councils premises, if applicable
- Cover remuneration, travel- and meeting expenses according to rules in the RCN
- Perform other necessary administrative functions related to the assignment
- Be responsible for contact between the Ministry and the evaluation committee
- Facilitate the dialogue between the evaluation committee and representatives of the advisory group and the involved scientific institutions. RCN shall be present at all meetings and should be notified of all contact between the evaluation committee and these groups
- Assure that the report is in accordance with the committee’s mandate
- Submit a completed report to the Ministry

Appendix B. The mandated reports included in this evaluation and their respective coverage

Type	Report	Index method	Panel method	Ecosystem
Framework	A technical system for assessing good ecological condition, 2017	X	X	All
Protocol	The Index-based Ecological Condition Assessment (IBECA) Technical Protocol version 1.0	X		
Protocol (peer reviewed article)	Introducing the Index-based Ecological Condition Framework (IBECA)	X		
Protocol	Panel-based Assessment of Ecosystem Condition (PAEC). Technical Protocol version 2		X	
Pilots	Test of the system for assessing ecological condition for terrestrial ecosystems in Trøndelag	X		Forest, mountains, wetlands, semi-natural land
Pilots	Panel-based Assessment of Ecosystem Condition – a pilot for four terrestrial ecosystems in Trøndelag		X	Forest, mountains, wetlands, open lowlands
First assessments	Norwegian Arctic tundra: a Panel-based assessment of ecosystem condition		X	Arctic tundra
First assessments	The condition of forest and mountain ecosystems in Norway: Assessment by the IBECA method	X		Forest, mountain
First assessments	Panel-based assessment of ecosystem condition of Norwegian Barents Sea Shelf Ecosystem		X	Ocean
First assessments	Panel-based Assessment of Ecosystem Condition of the North Sea Shelf Ecosystem		X	Ocean
First assessments	Panel-based Assessment of Ecosystem Condition of the Norwegian Sea Pelagic Ecosystem		X	Ocean

Appendix C. Interviews with stakeholders and experts

Interviews with stakeholders

The stakeholder interviews lasted for about an hour, were all held in Swedish/Norwegian by two evaluation panel members and then summarised into English for the entire evaluation panel to be informed. The 14 interviewed stakeholders represent the following:

- the Ministry of Climate and Environment (KLD),
- the Norwegian Environment Agency (MD),
- the Norwegian Agriculture Agency (LBD),
- the Directorate of Fisheries and the County Governor of Innlandet)
- the Norwegian Forest Owners' Association,
- World Wildlife Fund,
- Friends of the Earth Norway
- Sabima

Questions to stakeholders (free translation as the interviews were conducted in Swedish/Norwegian): *It should be noted that we conducted the interviews more as a conversation than a set of questions, and we posed follow-up questions as appropriate.*

- How well are you acquainted with IBECA/PAEC?
- Tell us more about how you have been in contact/involved with PAEC/IBECA!
- What do you think about the fact that there are two parallel systems?
- Which data sources/information systems do you have access to for assessing ecological condition of (as relevant for you) ecosystems?
- In what ways does this data/information come to use for your organisation?
- How to you value the usefulness of PAEC/IBECA for assessing ecological condition?
- To be useful for practice, how do you think the PAEC/IBECA results should be presented?
- Do you believe that the use of PAEC/IBECA will help to create consensus/legitimacy among different stakeholders? If so, how, and if not, why not?
- What challenges and opportunities do you see with the PAEC/IBECA methods in practice?

Interviews with assessment experts

Five assessment experts were heard to clarify questions we had about the methods. While two interviews with IBECA experts were conducted by two of the panel members (and thereafter translated to English), one session was carried out by the whole evaluation panel with three PAEC and IBECA experts together. Some remaining questions were later answered by email. The follow-up questions were too many questions to list here. The experts represent the following research institutions:

- Norwegian Institute for Nature Research in Trondheim, Tromsø and Oslo
- Institute of Marine Research
- University of Bergen

Norges forskningsråd
Besøksadresse: Drammensveien 288
Postboks 564
1327 Lysaker

Telefon: 22 03 70 00
Telefaks: 22 03 70 01

post@forskningsradet.no
www.forskningsradet.no

Publikasjonen kan lastes ned fra
www.forskningsradet.no/publikasjoner

Design: [design]
Foto/ill. omslagsside: [fotokreditt]

ISBN 978-82-12- 04189-9 (pdf)