Imperial College London

Department of Life Sciences

Effective design and use of indicators for marine conservation

Michael John Burgass

A thesis submitted for the degree of Doctor of Philosophy

Imperial College London

April 2019

Declaration of Originality

This thesis is a result of my own work. The work and contributions of others have been specifically indicated in the text.

Copyright Declaration

Abstract

The design, selection and use of indicators for large-scale conservation policy has been of great interest since the Convention on Biological Diversity (CBD) committed to a significant reduction in the rate of biodiversity loss by 2010. Following the introduction of the 2020 Aichi Targets, there was an increase, not only in demand for numbers of indicators, but the requirements that they are expected to meet. The complexities of social-ecological systems and the inevitable trade-offs that exist within them mean understanding and validating indicator responses are critical if they are to play a role in active management.

In this thesis, I look critically at uncertainties around how indicators are constructed and used, through the lens of marine science and conservation. I start the thesis by exploring the different types of uncertainty found when using composite indicators and from reviewing the literature, suggest possible methods of dealing with them. I find that structural uncertainties of indicators are rarely acknowledged. As a case study of application of composite indicators, I developed an Ocean Health Index assessment for the Arctic Ocean, demonstrating how a structured framework can be of great use for taking a data-driven approach to assessing social-ecological systems in large, data-poor regions. I show the Arctic is sustainably delivering a range of benefits to people, but with room for improvement in all areas, particularly tourism, fisheries, and protected places. Successful management of biological resources and short-term positive impacts on biodiversity in response to climate change underlie these high goal scores.

I then explore how two biodiversity indicators (Living Planet Index and Norway Nature Index) can be better interpreted and validated using an end-to-end ecosystem model, Atlantis, in the Nordic and Barents Seas. By simulating different fishing scenarios, I evaluated the extent to which the model-based testing approach gave insights into indicator behaviour; while the LPI is able to distinguish clearly between three different fishing scenarios, the NNI is only able to distinguish the most heavily fished scenario from the other two. I discuss how this approach

is useful for indicator testing and to advance integration of large-scale biodiversity indicators with goal-setting and decision making at the system scale. I then use the model to explore how different indicators of biodiversity from across fisheries and conservation respond to management interventions in Norway in the face of climate change. I find that despite having the same intentions, fisheries and conservation biodiversity indicators respond differently to each other under the same scenarios, due to how they are constructed. This means that without proper validation, indicators can potentially give different pictures of the same system to different interest groups, meaning greater integration and understanding of conservation and fisheries management objectives is necessary.

Finally, I reflect on the findings of my thesis in light of the CBD Post-2020 Framework. I discuss several core areas where the process could be revised to improve biodiversity outcomes. This includes formulating a robust theory of change to give the framework a clear conceptual basis and explicitly articulate the causal assumptions about the relationship between actions and outcomes. I do not focus on what targets should look like, but instead seek proactive, solutions-oriented approaches that can help 'bend the curve' for biodiversity.

This thesis highlights the uncertainties and challenges associated with large-scale indicator design and use and demonstrates how countries can take steps to reduce these. Greater consideration of the systems within which indicators are based can lead to better validation and ultimately better decision making.

Acknowledgements

There are many people who have played various roles in guiding me to this point. I must begin with a huge thank you to my most excellent supervisors, E.J. and Ben, who have been a source of great inspiration to me both professionally and personally over the course of the PhD. You both gave me great support and confidence throughout the PhD and encouraged me to pursue my other interests as well. Without this I would not have been able to be involved in such interesting side projects such as Conservation Optimism and more recently working with IIED to produce a handbook on Monitoring, Evaluation and Learning towards SDG 14. These experiences, as well as others, made the PhD into a sum of more than just these words and a truly life changing experience; which is exactly why I came back to do this. I thank you both for supporting me.

I would like to extend a huge thank you to all my other collaborators. Particularly to Emily Nicholson who has played a key role in stimulating a lot of thinking behind this thesis; your support over the years has been fantastic and I thank you for welcoming me to Melbourne, presenting with me at ICCB and your enthusiasm for the ICN workshop. A big shout out to all in your lab too, particularly Kate, Simone and Jess – I can't wait to read more of their work! Likewise, the OHI team at NCEAS – Julie, Jamie, Casey, Courtney and Mel. Thank you for your incredible kindness in welcoming me to Santa Barbara over the years and all the patience and help with R. You are the most awesome scientists and I feel lucky to be able to call you my friends. Hopefully I'll make it back over some day! Chapters 4 and 5 owe a huge thank you to Cecilie Hansen from the Institute of Marine Research in Bergen, Norway who showed great enthusiasm to share her modelling and work with me. Thank you also for being a brilliant host! Thanks also go to Bard Pedersen from the Norway Institute for Nature Research for being so open and willing to explore the NNI.

There are too many to name from ICCS as we have seen the group explode in numbers! But I feel very lucky to have been a part of such an amazing group and have learnt so much about so many topics, such is the diversity of people there. Thanks to University of Oxford who welcomed the Imperial College students with open arms and gave us a home.

Closer to home, in true Hollywood style, I must thank my mum, dad and wider family and friends for their continuous love and support. Without this and the stability you have given me, getting here would never have been possible. Thank you. Finally, to Katie, the true hero in all of this. Thank you for supporting the decision to do this and for your infinite patience. Your love, support and encouragement has kept me strong throughout. You embraced the move to Oxfordshire and I'm glad we were able to create so many lifelong memories there. This is not the end, but the new beginning I needed, and I can't wait to keep sharing it with you.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AOHI	Arctic Ocean Health Index
BII	Biodiversity Intactness Index
CBD	Convention on Biological Diversity
CCPI	Climate Change Performance Index
CI	Composite Indicator
DPSIR	Driver–Pressure–State–Impact–Response Framework
EBFM	Ecosystem Based Fisheries Management
EPI	Environmental Performance Index
FAO	Food and Agriculture Organisation
HLPF	High Level Political Forum
IOC	International Oceanographic Commission
LPI	Living Planet Index
MSE	Management Strategy Evaluation
MSP	Marine Spatial Planning
MSY	Maximum Sustainable Yield
NBSAP	National Biodiversity Strategy Action Plan
NNI	Norway Nature Index
OHI	Ocean Health Index
PREDICTS	Projecting Responses of Ecological Diversity In Changing Terrestrial Systems
RLI	Red List Index
SDG	Sustainable Development Goals
SSI	Sustainable Societies Index
ToC	Theory of Change
UN	United Nations
UNCLOS	United Nations Convention on the Law of the Sea
UNFCCC	UN Framework Convention on Climate Change
VNR	Voluntary National Review

1 INTRODUCTION

1.1 Background

Oceans cover 71% of the earth's surface and are home to up to 2.2 million species (Mora et al. 2017). The oceans regulate global climate, mediating temperature, driving weather systems and determining rainfall, droughts, and floods. Oceans also directly and indirectly contribute to the wellbeing of society: Most of the world's megacities are located in coastal zones and coastal communities often have deep-rooted interconnections with the ocean culturally, socially and physically (von Glasow et al. 2013). Around 3 billion people source nearly 20% of their mean daily animal protein intake from the oceans, providing nutritional and health benefits, which are crucial for poverty and hunger reduction (Bailey et al. 2016). However, the interaction between people and the environment has meant that global biodiversity is threatened by a range of pressures including over-exploitation of species, habitat modification, invasive alien species and disease, pollution, and climate change (Maxwell et al. 2016). In marine systems, such pressures stem from a range of activities such as fishing, coastal development, shipping and energy production (Halpern et al. 2015a), which has left only 13% of global ocean as 'wilderness' (Jones et al. 2018a). Managing marine systems is therefore a complex endeavour, which ideally would result in 'triple bottom line' outcomes, where conservation goals and social outcomes are maximised and overall costs are minimised (Halpern et al. 2013a).

1.1.1 Integrated Marine Management

An integrated management approach, considering environmental, social and economic outcomes, changes the paradigm of traditional marine management, where individual pressures are managed separately, to managing activities in combination through a holistic ecosystem-based management approach (Borja et al. 2016). However, doing so is not straightforward as each sector tends to have its own organisational bodies, systems, frameworks and priorities which are a barrier towards true integration (Elliott 2014). Nevertheless, a range of different systems and tools have been developed for managing marine systems in an attempt to balance economic and social development with environmental protection.

In reality this means that fisheries management should consider conflicting objectives in decision-making, but instead typically tends to focus on minimising impacts to aspects such as threatened species and habitats, through by-catch or gear usage (Milner-Gulland et al. 2018). The ecosystem approach to fisheries management or ecosystem-based fisheries management (EBFM) is a term that was formally accepted at the Earth Summit in Rio de Janeiro in 1992. The United Nations Food and Agriculture Organisation (FAO) states that "An ecosystem approach to fisheries strives to balance diverse societal objectives, by taking into account the knowledge and uncertainties about biotic, abiotic and human components of ecosystems and their interactions, and applying an integrated approach to fisheries within ecologically meaningful boundaries" (FAO 2003). Fish stock productivity, and thereby sensitivity to harvesting, depends on physical (e.g. ocean climate) and biological (e.g. prey availability, competition and predation) processes in the ecosystem (Serpetti et al. 2017). While traditional fisheries management focuses on harvest rates and stock biomass, incorporating the impacts of such ecosystem processes on fish stocks is one of the main pillars of EBFM. However, despite much attention in the literature, EBFM is yet to be widely implemented (Skern-Mauritzen et al. 2016). Nonetheless, advances in decision-making theory and practice through processes such as Management Strategy Evaluation (MSE) have been shown to enable fisheries managers to balance conflicting priorities in marine successfully systems (Fulton et al. 2014).

Marine Spatial Planning (MSP) has also proved a popular as a spatial tool for controlling marine development. MSP is a process of analysing and allocating the spatial and temporal distribution of human activities in marine areas to achieve ecological, economic, and social objectives that usually have been specified through a political process (Agardy et al. 2011). It is now used in many countries across the world and has developed a strong community of practice and guidance through the International Oceanographic Commission of UNESCO (IOC-UNESCO) to assist with new implementation. MSP and EBFM tend to operate separately, despite both having intentions of being truly integrative, although it is possible to implement MSP with minimal loss to fisheries yield (Klein et al. 2010).

1.1.2 International agreements

These marine management approaches operate within a complicated legislative framework, which covers multiple sectors at multiple scales from local to global. Marine systems by their nature are dynamic and cross-boundary and as such international agreements are necessary for ensuring coordinated management (Molenaar 2015). In 2015, the United Nations (UN) agreed on 169 targets to mobilize action towards sustainable development. The Sustainable Development Goals (SDGs) are 17 broad goals covering a broad range of areas from poverty reduction, to clean energy and economic growth. Among those goals, Goal 14: Life Below Water aims to "conserve and sustainably use the oceans, seas and marine resources for sustainable development". The targets underpinning this goal range from marine pollution, to fisheries management and knowledge transfer. SDG 14, along with SDG 15 (Life on Land) are the two SDGs which enshrine ecosystem and biodiversity protection within them. These were important goals for the conservation sector as biodiversity has been decreasing at such a rate that in 2002, through the Convention on Biological Diversity (CBD), 196 countries committed to "to achieve by 2010 a significant reduction of the current rate of biodiversity loss." In 2011, the ambition of the CBD was strengthened, as the Aichi Targets were implemented, and many of these were integrated into the SDGs with the overall aim of SDG 14 to halt biodiversity loss by 2030 (United Nations 2015). Aichi Target 6 specifically aims to sustainably manage marine resources using ecosystem based approaches.

Both the SDGs and the Aichi Targets explicitly allow countries to determine their own conservation strategies, considering the very different contexts between countries (Singh et al. 2018). The SDGs coordinate this around the High-Level Political Forum (HLPF). The HLPF is a voluntary and state-led body, which provides a platform for partnerships, including through the participation of major groups and other relevant stakeholders. It has the responsibility for tracking SDG implementation and achievement through Voluntary National Reviews (VNRs). The 2030 Agenda encourages member states to "conduct regular and inclusive reviews of progress at the national and sub-national levels, which are country-led and country-driven". These VNRs serve as a basis for the regular reviews by the HLPF of progress for each SDG (Committee for Development Policy 2018). Similarly, the CBD requires members both to produce a plan for achieving the Aichi Targets through National Biodiversity Strategy Action Plans (NBSAPs), but then also to report on progress towards success through the National Reports (Secretariat of the Convention on Biological Diversity 2016).

Outside of the CBD and SDGs, many other international agreements apply to the marine realm. Of particular importance is the United Nations Convention on the Law of the Sea (UNCLOS), an international agreement signed in 1982 that sets out the legal framework for all activities in the oceans and seas and is of strategic importance as the basis for cooperation in the marine sector. UNCLOS has 168 parties, including the EU, but is considered in most of its provision as customary international law (Hoagland et al. 2001). It was included as a key factor in SDG 14 by being included in Target 14.C "Enhance the conservation and sustainable use of oceans and their resources by implementing international law as reflected in UNCLOS, which provides the legal framework for the conservation and sustainable use of oceans and their resources, as recalled in paragraph 158 of The Future We Want". It is complemented by the 1995 UN Straddling Fish Stocks Agreement, which sets out principles for the conservation and management of fish stocks and establishes that their management must be based on the precautionary approach and best available scientific information. The Agreement elaborates on the fundamental principle of UNCLOS that States should cooperate to ensure conservation and promote optimum utilisation of fisheries resources. These ideals are implemented through a framework known as the Code of Conduct for Responsible Fisheries (FAO 1995).

1.1.3 The role of indicators

Assessment of progress towards national and international environmental ambitions is typically monitored through a series of indicators (Hák et al. 2016a). Monitoring is important for three main reasons; to inform decison-makers when a system is departing from a desired state, to measure the success of management actions and to detect the effects of disturbance or stressors (Legg & Nagy 2006). Indicators often play a major role in monitoring, as they allow for the communication of complex information in a more simplistic manner to a range of stakeholders and sectors such as governments and business (Jørgensen et al. 2013). Criteria for indicator selection are widely discussed in the literature and are generally agreed to include aspects such as measurability, scientific basis, ease of communication, sensitivity and responsiveness to change, and specificity (Failing & Gregory 2003; Rice & Rochet 2005; Niemeijer & de Groot 2008a). Increasingly there is greater understanding and importance placed upon interpretability, validation and in turn understanding trigger points or thresholds for management interventions (Samhouri et al. 2010; Large et al. 2013; Moriarty et al. 2018).

A key role of indicators is at the nexus of science and policy, as their purpose is often to communicate with broad and non-specialist audiences, requiring scientific rigour but straightforward communication. Typically science and policy has had a complex relationship as science is rarely used as a sole guide for decision making (Bradshaw & Borchers 2000). However, indicators are often primarily designed by scientists as scientific or communication tools, with little attention given to the political landscape (Robertson & Hull 2001). A recent review of species indicators found that only 21% explicitly accounted for management objectives and actions(Bal et al. 2018). For science to be useful to policy makers it must be perceived as credible, salient and legitimate (McNie 2007). Credibility refers to the scientific adequacy of the technical evidence and arguments. Bradshaw & Borchers (2000) note that uncertainty is a key blockade in bridging the science-policy gap and can undermine credibility. They note that although scientists are comfortable with uncertainty, decision makers are on a quest for certainty and deterministic solutions. Salience refers to the information being relevant to the specific context in which it will be used and responsive to the specific information demands of the decision makers. Legitimacy reflects the perception that the production of information and technology has been respectful of stakeholders' divergent values and beliefs, unbiased in its conduct, and fair in its treatment of opposing views and interests (Cash et al. 2003). All three of these are heavily interlinked such that efforts to enhance any one of these usually incur a cost to the others, meaning a successful balance must be struck (Cash et al. 2003).

1.1.4 Challenges with indicators

A challenge in developing large-scales indicators is that they have either been developed by scientists without broader consideration of the use of the indicator in the political/management process, or in the case of international agreements, targets have been set without consideration of what indicators can be measured (Turnhout et al. 2007; Maxwell et al. 2015). This means that it is challenging to develop appropriate indicators. In the years after the CBD's goal to achieve a reduction in the rate of biodiversity loss had been agreed, a significant effort was made to design and select indicators to monitor progress towards this goal (Mace & Baillie 2007). Nonetheless, by 2010 the indicator set was still not complete (Walpole et al. 2009). A similar problem was later identified for the 2011-2020 Aichi Targets, such that again these targets appear unlikely to be met (Tittensor et al. 2014). Recognising the intrinsic links between biodiversity and human wellbeing, in 2002 the Aichi Targets, as well as the SDGs, included more human-centric targets related to concepts such as ecosystem services (Shepherd et al. 2016). However, such targets have received criticism for being overly complex and ambiguously worded (Butchart et al. 2016). This has meant that indicators are poorly aligned towards their targets (Mcowen et al. 2016). To be most effective, indicators should be created in the domains of both science and policy and go back and forth between them until there is consensus (Turnhout et al. 2007). However, in reality targets are heavily negotiated in political forums, so that this is not easy to achieve.

The CBD process for the Aichi Targets has assigned multiple indicators to each individual target (Convention on Biological Diversity 2016). This is done by breaking down each target into sub-components and then assigning an indicator or indicators to each sub-component, based on a series of criteria. This differs somewhat to the SDGs, which nominate a single indicator for each sub-target of each SDG. This theoretically makes assessment more straight forward, but in reality suffers issues of indicator and target alignment, particularly when the sub-targets are multi-faceted (Mcowen et al. 2016). Multiple indicators are often required but this presents challenges regarding distilling information (Chatziparadeisis 2007). Dashboards have found some use in displaying such information (Han et al. 2014), but composite measures are finding increasing use for their ability to measure different types of indicators on the same scale (Munda et al. 2009).

A key issue regarding the science-policy interface of indicators is around scale. In the CBD and SDGs, indicators are assigned to global targets, but individual nations are also required to report their national progress using indicators through the voluntary processes of the National Reports and VNRs (Hagerman & Pelai 2016). It has therefore been a priority to ensure that biodiversity indicators are applicable across scales. This has largely resulted in indicators containing a spatial element, which allows them to be disaggregated at different scales. For instance, the Living Planet Index (LPI) records locations and biomes of where each data point is recorded so that it can be later disaggregated (Mcrae et al. 2012). The Ocean Health Index (OHI) likewise has the ability to be disaggregated, but has also been designed so that the framework can be reused at different scales and global data can be replaced with more relevant local data (Halpern et al. 2012; Elfes et al. 2014). However, the extent to which CBD indicators are used at the national scale is largely unknown.

While ensuring indicators are useful to end-users, they must also be scientifically credible. An indicator's usefulness can be impeded by both the quality of the data which underpins it and the design of the indicator itself (Collen & Nicholson 2014). Indicators tend to utilise secondary data, meaning the quality of data in theory should be assessed and any uncertainties understood and estimated (Regan et al. 2002; Munda et al. 2009). In reality, this is often not the case as data quality is traded-off with availability, particularly where monitoring data is scarce (Griffiths et al. 2010). This is particularly an issue for large-scale indicators, which utilise freely or widely available data. For biodiversity indicators, this might mean geographic or taxonomic biases reduce an indicator's overall representativeness of the biodiversity it is attempting to measure (Nicholson et al. 2012). The way an indicator is constructed, particularly with aggregate indicators, may make its validation and interpretation challenging, thus impacting its credibility (Moriarty et al. 2018). How indicators are weighted has been widely discussed, particularly for composite indices, as it can highly impact the outputs of indicators and often has a weak basis (Saisana et al. 2005). While indicator construction is widely discussed, there are few examples of conservation indicators being formally tested and validated, despite this being a priority under the CBD (Collen & Nicholson 2014).

In fisheries science, indicators have been the focus of widespread attention to formally evaluate their usefulness in detecting the effects of changing fishing pressure (Fulton et al. 2005; Shin et al. 2018). Such approaches are extremely important in helping to understand indicator behaviour and thus in directing robust management interventions (Fay et al. 2013). Gaining such understanding of large-scale indicators is rarely undertaken and is needed given the complex social-ecological systems within which people and biodiversity co-exist (Hill et al. 2016). These complex systems provide a significant challenge for traditional target development and indicator selection approaches, as feedbacks and interactions can change parts or all of the system in unforeseen ways (Ostrom 2009; McGinnis & Ostrom 2014; Larrosa et al. 2016). Despite being of central importance, these interlinkages are rarely considered up front, as targets are negotiated politically (Maxwell et al. 2015). This has meant that there are trade-offs and synergies within the targets which prevents targets simply being ‘ticked off’ (Di Marco et al. 2016; Singh et al. 2018). It is important, therefore, that changes in indicators are understood within the context of the wider social-ecological system they are within, such that they can be appropriately interpreted and used correctly.

1.1.5 Modelling and scenarios

When it comes to implementing policy to meet high level political targets, the Driver–Pressure–State–Impact–Response (DPSIR) framework, and variants of it, have proved popular for emphasising the importance of causality (Gari et al. 2015). Although it structures and standardizes conceptualizing complex issues, it has been criticised for providing an overly simplistic representation of the relationship between pressures and state changes, by assuming that increases in pressures lead to state changes, which may not always be the case (Smith et al. 2016). It is unable to take account of the interactions between different activities and their cumulative pressures occurring simultaneously, which we know to be important, particularly in the marine environment (Halpern et al. 2015a; Patrício et al. 2016). Furthermore, being a simple unidirectional chain it does not highlight the difference in the nature, severity, timescale or longevity of state changes in relation to pressure intensity, frequency or duration and thus is not particularly conducive to representing an understanding of the complexity of the processes within systems and thus behind environmental indicators (Niemeijer & de Groot 2008b). Expanding such frameworks out to network approaches is seen as a useful way of improving indicator selection and use (Niemeijer & de Groot 2008a; Vugteveen et al. 2015).

The use of modelling and scenarios has been proposed as a key way of understanding systems, setting science-based targets and informing appropriate indicator selection and use in conservation science (Nicholson et al. 2019). Models are simplified understandings of systems and range from qualitative conceptual models, which display linkages and relationships between different elements of a system (Margoluis et al. 2009), through to quantitative models which are built from an in-principle understanding of the system or from analysis of emergent patterns of data (Jørgensen 2008). Scenarios depict different future states, often influenced through different management interventions. Models and scenarios are typically used together to explore management options and have been shown to be useful in conservation for testing indicator responses. Costelloe et al. (2015) demonstrated how both the Red List Index (RLI) and LPI indicated more effective management would provide greater benefits to biodiversity than merely expanding protected areas, when modelling management options for Sub-Saharan African protected areas. Nicholson et al. (2012) showed how taxonomic bias in the LPI made its interpretation difficult when simulating the effect of a marine management intervention (an end to bottom trawling), while Visconti et al. (2016) demonstrated business as usual development approaches would mean Aichi Target 12 (improving conservation status of known threatened status) cannot be achieved. The Biodiversity Intactness Index (BII) is currently being used with the Projecting Responses of Ecological Diversity In Changing Terrestrial Systems (PREDICTS) project, using annual fine-scale pressure data to drive annual estimates of how BII has changed in the recent past and also using modelling to project the BII using historical and future estimates of land use and other pressures from the Shared Socioeconomic Pathways (Scholes & Biggs 2005; Mace et al. 2014; Purvis et al. 2018).

1.2 **Aims & Objectives**

The aim of this thesis is to investigate the challenges related to the use of large-scale environmental indicators and explore how these can be addressed. I frame this investigation through the lens of marine conservation, using marine systems and indicators as case studies. This is with the aim of informing the creation and use of such indicators for the marine environment, but more generally in wider national and international conservation.

Specific objectives that contribute to the aim of the thesis are to:

- Summarise uncertainties associated with composite indicators and review methods of treating them;
- Explore composite indicator usage in a data-poor context;
- Apply a modelling approach to indicator validation at the national scale;
- Explore future scenarios to compare biodiversity indicators across fisheries and marine conservation;
- Explore applications of environmental indicators in future international agreements.

1.3 **Thesis Structure**

Figure 1-1 below shows a conceptualisation of the thesis structure. It starts with understanding the background to the topic (Chapters 1 & 2), before exploring two approaches to indicator construction and use; a structured framework approach (Chapter 3) and a systems-based approach (Chapter 4). Elements of these two approaches are integrated and compared in Chapter 5, before Chapters 6 and 7 explore the wider context of international indicator development and use, and application of the thesis's insights in the future.

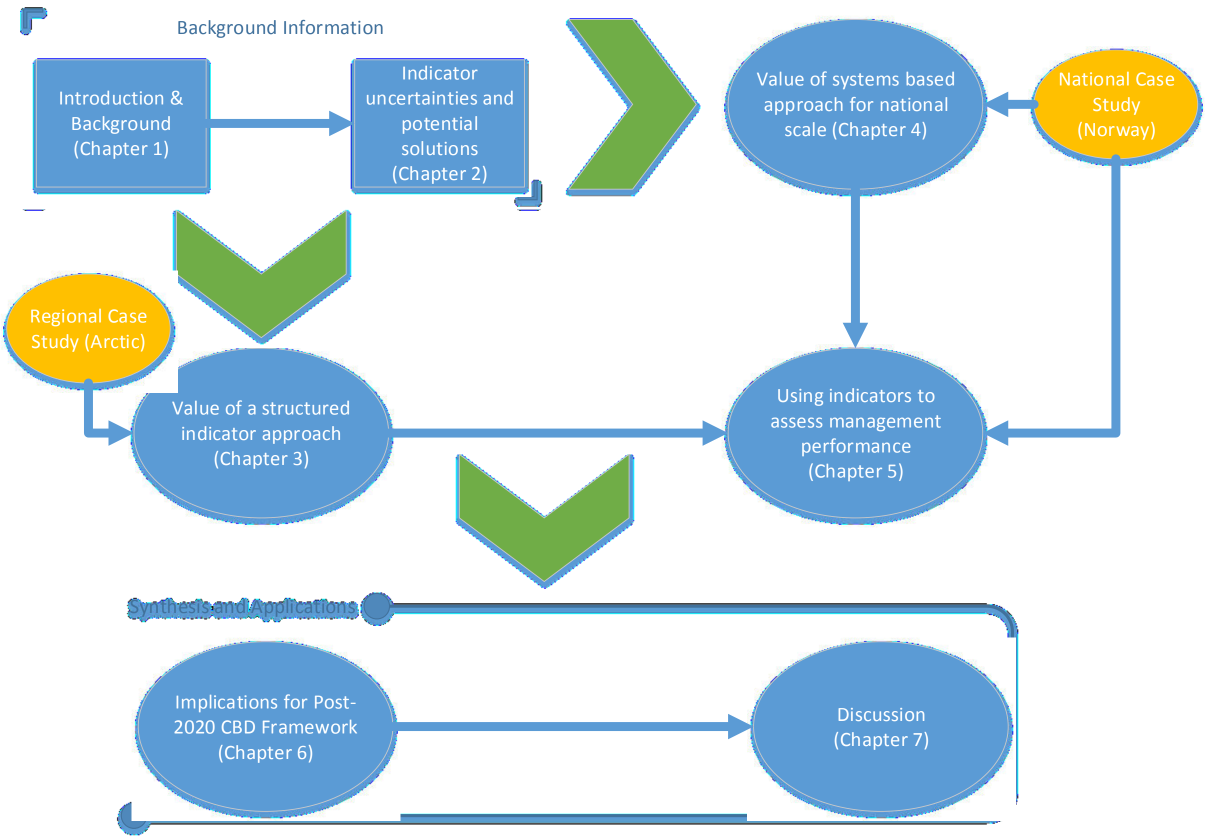

Figure 1-1: Conceptualisation of thesis structure

1.4 Case Study Selection and Overview

In Chapters 3, 4 and 5 I use two case studies to explore indicator design and use. These were purposefully selected to explore the need for both structured and systems-based indicators at different scales. Chapter 3 takes a structured approach to the Arctic region, where data are often lacking, and management approaches differ. Chapters 4 and 5 zoom in to the Barents and Nordic Seas, primarily from a Norwegian perspective. Here I focus on large-scale biodiversity indicators and how they are used at the system level.

1.4.1 Case Study 1 – The Arctic Ocean

The Arctic Ocean is the smallest and shallowest of the world's five oceans and unique in that it is mostly enclosed with limited exchange of water with other seas or oceans. It has a complex circulatory system which drives the accumulation and melt of sea ice. Sea ice is the key driver of life in the Arctic Ocean, supporting vast amounts of algal and phytoplankton primary production (Arrigo et al. 2012) which provide the basis for the Arctic food web. Life here endures some of the greatest extremes in light and temperature known to our planet, but the Arctic Ocean contains a rich tapestry of benthos, fish, cnidarians, birds, and some of the most recognisable arctic mammals such as cetaceans, pinnipeds and of course polar bears (*Ursus maritimus*).

The Arctic is also responsible for regulating global climate as sea ice acts as an important constraint on methane release from clathrates and permafrost (Parmentier et al. 2013) as well as making a significant contribution to the earth's surface albedo (Deser et al. 2000). Sea ice also plays a role in thermohaline circulation as water cools as it enters the Arctic. Freezing water rejects its salt content causing surface waters to increase in salinity, sink to the bottom and flow out again. This process of North Atlantic Deep Water Formation is critical to global thermohaline circulation (Dickson & Brown 1994).

The Arctic Ocean's biophysical processes are therefore extremely important for regulating global climate and supporting fragile ecosystems, but when also considering human interests and activities in the Arctic and its sensitivity to change, it becomes one of the most geo-politically important areas on the planet. The food web supports globally significant fisheries of pollock (*Theragra chalcogramma*) and cod (*Gadus morhua*), the seafloor is thought to contain approximately 13% and 30% of the world's undiscovered oil and gas reserves respectively (Bird et al. 2008) and shipping through the Northern Sea Route is increasing annually (Northern Sea Route Information Office 2013), with future projections signalling exponential escalation (Smith & Stephenson 2013). Furthermore the Arctic is inhabited by roughly four million people of whom approximately 400,000 are indigenous (Arctic Council 2011). Climate change is beginning to affect the Arctic Ocean with observations of range shifts and changes in abundance, growth and phenology of species (Wassmann et al. 2011). Sea ice reduction is also leading to the rush for new shipping routes, mineral resources and fishing grounds that have previously been inaccessible and security and defence is also becoming a growing issue, with some Arctic states bolstering their forces in their northernmost territories (Kraska 2011).

The issues raised above are cross-sectoral and dynamic in nature, meaning management interventions or changes in practices can affect several components of the Arctic Ocean system i.e. implementation of protected areas could cause economic losses from lower fish catches, ecosystem shifts, increases in tourism revenue etc. However, research in the Arctic is usually specific to an individual country or focussed on one dimension of the system (biological, physical or social), despite the role of integration being well understood in achieving successful conservation and coastal management (Turner 2000; Ban et al. 2013). Calls for greater unity in Arctic research are not new (UNEP/GRID Arendal 2006) but existing marine monitoring efforts are still not connected on a circumpolar scale, which limits the ability to make robust decisions. It would appear Arctic Ocean research is relatively disconnected and only now are measures to standardise data collection and outputs being organised (i.e. Circumpolar Biodiversity Monitoring Program). Given that the Arctic is potentially changing rapidly and data poor, it is an exciting case study to explore the value of a structured indicator approach to compiling data and attempting to measure both the biophysical and socio-economic dimensions of the ocean ecosystem.

1.4.2 Case Study 2 – The Nordic and Barents Seas

Norway has a long Arctic and sub-Arctic coastline which borders the Nordic and Barents Seas and as such fisheries have been important culturally and socially important for livelihoods in Norwegian coastal communities for centuries. In the late 1960s Norway experienced dramatic effects on fisheries and coastal communities due to the collapse of the large Norwegian spring spawning herring stock, caused by overfishing. The continuation of overfishing also had detrimental effects on the large fish stocks in the Barents Sea, the Norwegian Sea, and the

North Sea following rapid technical progress and increased efficiency in the fisheries (Gullestad et al. 2014). This gave the Norwegian government the stimulus to reform the fishing sector, which it did so over the next 40 years through limiting access, ending subsidies, reducing overcapacity, improving quota distribution, altering discard strategy, implementing a precautionary approach and improving co-management (Petter Johnsen & Eliasen 2011; Gullestad et al. 2014; Grønnevet 2016). The fisheries management system in Norway is now rated among the highest anywhere in the world for compliance to UN Code of Conduct for Responsible Fisheries (FAO 1995), in which prevention of overfishing is among the central principles (Pitcher et al. 2009a, 2009b). Nonetheless, overall compliance to the Code of Conduct by Norway was about 60%, still indicating considerable potential for improvement.

In 2009, a new Marine Resources Act entered into force in Norway. The act shifted the focus from merely managing commercial exploitation of marine resources, to all wild living marine resources and genetic material derived from them (Gullestad et al. 2017). The act states that its purpose is to ensure sustainable and economically profitable management of the resources, and several provisions describe conservation of biodiversity as an integral part of sustainable management, including requiring “an ecosystem approach, taking into account habitats and biodiversity” (Norwegian Government 2010). As part of Norway’s shift towards establishing an ecosystem approach, they have implemented a large-scale habitat mapping project, known as MAREANO (Buhl-Mortensen et al. 2015) and developed an end-to-end ecosystem model, Atlantis, for the Nordic and Barents Seas (Hansen et al. 2016, 2019). This ecosystem model and the associated thinking provided an exciting opportunity to explore and demonstrate a systems-based approach towards indicators.

1.5 Thesis Outline

1.5.1 Chapter 2: Navigating uncertainty in environmental composite indicators

In this chapter I review the different sources and types of uncertainty found within composite indicators, from initial construction to communication and use. I review the literature to suggest key areas for how to reduce this uncertainty and use four well known composite indicators as case studies to analyse how well this is being done in practice. I find that while there are many uncertainties, there are many different potential methods for treating uncertainty. In general these are poorly implemented for current indicators.

This chapter has been published as:

Burgass, M.J., Halpern, B.S., Nicholson, E. & Milner-Gulland, E.J. (2017). Navigating uncertainty in environmental composite indicators. *Ecological Indicators*, 75, 268–278.

I conceived the research idea with E.J. Milner-Gulland and Ben Halpern. I chose case-study indicators, conducted the review and wrote the paper. All co-authors provided comments and revisions.

1.5.2 Chapter 3: A pan-Arctic assessment of the status of marine social-ecological systems

In Chapter 3 I demonstrate how the structured composite indicator approach can be useful for assessments across large, data-poor areas. I use the Ocean Health Index across the Arctic Ocean to consolidate data for nine different Arctic regions across nine areas of ocean health. I make the first assessment which considers social-ecological conditions across the Arctic. I use two different data sets for the fisheries sub-goal, to show how underlying data choices can impact scores.

This chapter has been published as:

Burgass, M.J., Milner-Gulland, E.J., Stewart Lowndes, J.S., O'Hara, C., Afflerbach, J.C. & Halpern, B.S. (2019). A pan-Arctic assessment of the status of marine social-ecological systems. *Regional Environmental Change*, 19, 293–308.

I conceived the research idea with Ben Halpern. I identified regions and datasets, wrangled data, modified goal models, developed code and wrote the manuscript. Julie Stewart Lowndes, Jamie Afflerbach and Casey O'Hara provided technical assistance on code development and visualisation. All co-authors provided comments and revisions.

1.5.3 Chapter 4: Validation and use of large-scale biodiversity indicators at the national scale

Validation of global species-based indicators to test their usefulness at the national scale, while being a priority, has yet to be widely undertaken. In Chapter 4, I demonstrate how systems thinking can support indicator validation at the national scale in the marine environment. I show how two different indicators, the global Living Planet Index and national Norway Nature Index, can be generated from ecosystem models and their performance explored under different management scenarios. Widespread uptake of such a validation approach could help with developing more robust indicators, more meaningful projections of biodiversity change into the future and explicit and science-based target-setting at a range of scales, by allowing a better understanding of how indicators perform under different scenarios.

This Chapter was developed with the following co-authors: E.J. Milner-Gulland, Ben Halpern, Emily Nicholson, Cecilie Hansen and Bård Pedersen. I conceived the study with E.J. Milner-Gulland, Ben Halpern and Emily Nicholson. Cecilie Hansen provided ecosystem modelling. I analysed model outputs, developed code, generated indicators, conducted analysis and wrote the manuscript. Bård Pedersen assisted with coding for the Norway Nature Index. All co-authors provided comments and revisions.

1.5.4 Chapter 5: Assessing biodiversity loss with fisheries and conservation indicators

In this Chapter I apply three contrasting fisheries management approaches within a systems model for the region of the Nordic and Barents Seas over a period of 38 years, simulating forwards from the present day under a climate change scenario. I do this with two aims; to see if these management options could halt and potentially reverse biodiversity loss in line with international commitments in the face of climate change, and to see how consistent these predictions were between indicator types. I find that fisheries and conservation indicators disagree on whether biodiversity loss is halted as a result of the management changes, due to how they are constructed and what they intend to measure. In Norway, fisheries are the dominant sector in which biodiversity is managed, but the types of indicators used to do this are not necessarily well aligned with conservation objectives. This is shown to be problematic as fisheries ecosystem indicators do not always reflect the wider ecosystem, particularly aspects of conservation concern.

This Chapter was developed with E.J. Milner-Gulland, Cecilie Hansen and Ben Halpern. I conceived the study with E.J. Milner-Gulland. Cecilie Hansen provided ecosystem modelling. I analysed model outputs, developed code, generated indicators, conducted analysis and wrote the manuscript. E.J. Milner-Gulland and Ben Halpern provided comment and revision.

1.5.5 Chapter 6: Opportunities for setting a successful post-2020 global biodiversity framework

In Chapter 6, I reflect on key considerations for a solutions-oriented approach to setting a post-2020 framework to improve outcomes for biodiversity. Three core areas that a post-2020 framework must consider are: 1) Formulating a robust Theory of Change to link outcomes and actions; 2) Being underpinned by models to integrate complexity and uncertainty; and 3) Transcending scale to inform meaningful devolved and specific local action. With reference to each consideration, I discuss opportunities for improving the processes around how global targets are set and implemented by drawing on a range of examples.

This Chapter is the output of a three-day workshop was held in Oxford in July 2018 bringing together academics and conservation practitioners to share lessons learnt and discuss ways forward for international biodiversity commitments. I planned the workshop with Siso Larrosa and E.J. Milner-Gulland. Siso Larrosa and I ran the workshop. All participants contributed to conception of the paper and provided text. I led the development of the paper following the workshop including drafting the outline, consolidating text, writing the manuscript and organising and responding to co-authors. The list of co-authors not yet mentioned is as follows: Derek Tittensor, Emily Nicholson, Kate Watermayer, Jessica Rowland, Victor Muposhi, Shannon Hampton, Hernan Caceres, Abbey Camaclang, Carolina Pinto, Ciaran McLaverty and Simone Stevenson.

1.5.6 Chapter 7: Synthesis and Discussion

I conclude the thesis by summarising the main findings and discussing their relevance for informing future indicator development and use at large scales. Recommendations are given for how the lessons learnt from the thesis can be applied in practice.

2 NAVIGATING UNCERTAINTY IN ENVIRONMENTAL COMPOSITE INDICATORS

2.1 Introduction

Human activities have large impacts on natural systems (Halpern *et al.* 2008; Buma & Wessman 2011) that are likely to increase in future, given growing human population and demand on natural resources (Kraxner et al. 2013; McCauley et al. 2015). The resultant changes in natural systems have important consequences for biodiversity (Chapin et al. 2000), but also for people through our reliance on provision of ecosystem services for human well-being, health, livelihoods and survival (Costanza et al. 1997, 2014; Millennium Ecosystem Assessment 2005). Managing these complex interactions to ensure nature thrives and continues to provide benefits to people requires integrative and interdisciplinary approaches to management that emphasise the complexities of whole social-ecological systems (Folke et al. 2005). Effective ecosystem management requires measuring the status and trends of ecosystems to inform which management actions are likely to be effective and if these actions have had their intended effect (Jones et al. 2011). Measuring all aspects of complex systems is impossible due to the range of variables and processes present. Variables deemed to be characteristic of the wider system and which are simple enough to be easily measured are often employed as indicators, to act as simplified summaries of system condition and behaviour (Dale & Beyeler 2001).

Good indicator design has been widely discussed (Failing & Gregory 2003; Fulton et al. 2005; Parr et al. 2010), with general agreement that indicators should: be cost effective; provide reliable information on status and trends; provide information at multiple extents and resolutions; allow frequent reporting; be meaningful to the public; and respond predictably to policy change (Jones et al. 2011). In practice, the EU's Streamlining European Biodiversity Indicators project used a stakeholder-based process to apply stringent criteria and reduce over 140 biodiversity indicators to a final 26, while the European Commission assesses indicators based on RACER guidelines; where they should be ‘Relevant’, ‘Accepted’, ‘Credible’, ‘Easy to Evaluate’ and ‘Robust’ (Best et al. 2008; Eea 2010). Indicators also provide a powerful tool for communicating with stakeholders about the status and trends of ecosystems, as well as helping identify or illuminate linkages between environmental, human and economic subsystems (Jørgensen et al. 2013). However, multi-dimensional processes (such as complex ecosystem dynamics) are notably difficult to track with individual indicators due to challenges in linking trends across dimensions (Munda 2005) and capturing interactions between and within sub-systems (Dale & Beyeler 2001). Multiple indicators are recommended to capture different aspects of the relevant systems (Fulton et al. 2005), but without techniques to distil or summarise them, can be overwhelming in volume of information (Chatziparadeisis 2007). For example, the marine “Good Environmental Status” goal for EU countries contains 11 descriptors with 29 criteria and 62 individual indicators (European Comission 2010).

“Composite indicators” (CIs) offer a means of aggregating multiple indicators to track and communicate complex systems. CIs are a mathematical combination of a set of indicators that have no common meaningful unit of measurement. They are increasingly used for decision making in a range of sectors such as economics, business statistics, health and academic performance (Munda et al. 2009; Paruolo et al. 2013). In the environmental sector they are often used for global scale assessments (see Table 2-1) and to guide policy at local to regional scales (Mendoza & Prabhu 2003; Di Franco *et al.* 2009; Ochoa-Gaona *et al.* 2010). CIs enable direct comparison of disparate social and environmental variables and, due to their clear and unidimensional output, can also gain traction with policy-makers and the general public. Their increasing popularity is unlikely to slow; many have suggested that in order to communicate broad trends effectively and influence conservation policy, meaningful CIs will be required (Balmford *et al.* 2005; Mace & Baillie 2007).CIs are similar to mathematical or computational models in that they are simplified representations of reality, although whereas models are usually based upon scientific theory and detailed biological or physical dynamics, CIs are often simply an aggregation of variables considered relevant to a system or issue

(Nardo *et al.* 2008). Modelling studies also typically address, and when possible quantify, inherent uncertainties that arise when simplifying real-world complexities (Kokko 2005). If CIs are to be used more, and more effectively, within conservation, methodological decisions made in their construction, and the consequent uncertainties, should be clearly understood, described and, if possible, represented or treated – just as with any other type of conservation modelling for decision-making (e.g., Regan, *et al.* 2002).

Here, I explore the uncertainties that underlie environmental CI construction, with the aim of putting recognition of uncertainty at the heart of CI construction and use. I develop a framework to capture the full range of types and sources of uncertainty in a systematic fashion, using four prominent environmental CIs as primary case studies (but also draw reference to others) and suggest methods to navigate them. I first discuss the methods that are specific to each individual stage and then address those that deal with multiple sources of uncertainty. Finally, I discuss ways forward to improve the development and use of composite indicators in practice.

Table 2-1 Examples of environmental composite indicators, chosen to display a range of different construction techniques

Composite Indicator	**Description**	**Construction**
Ocean Health Index (Halpern *et al.* 2012) www.oceanhealthindex.org	Evaluates the condition of marine ecosystems according to ten ‘goals’ of key benefits provided by the ocean. Measures sustainable provision of benefits and gives a score to each country.	Overall score is aggregated from ten equally weighted categories (known as ‘goals’, each comprised of many individual indicators. Subgroups measure biological, physical, social and economic aspects.
Environmental Performance Index (Hsu *et al.* 2014) www.epi.yale.edu/	Ranks how well countries perform on high priority environmental issues. Focuses on ranking individual countries.	Nested structure where overall score is aggregated from two equally weighted categories of environmental health and ecosystem vitality. Each category is made up of three and six subgroups respectively, which have between one and four sub-indicators each.

Table 2-1 Examples of environmental composite indicators, chosen to display a range of different construction techniques

Composite Indicator	Description	Construction
Climate Change Performance Index (Burck & Bals 2011) www.germanwatch.org/en/9472	Evaluates and compares the climate protection performance of countries that are, together, responsible for more than 90% of global energy-related CO_2 emissions.	Nested structure where overall score is aggregated from three categories of emissions trend (50% weighting), emissions level (30% weighting) and climate policy (20% weighting). Each category is made up of between 4 and 9 subgroups which are informed by several sub-indicators each
Sustainable Society Index (van de Kerk *et al.* 2014) http://www.ssfindex.com/	Evaluates countries based on their level of sustainability according to human, environmental and economic wellbeing. Focusses on ranking of countries.	Employs a nested structure with three categories; human wellbeing, economic wellbeing and environmental wellbeing. Categories are not aggregated to an overall score due to the correlation between human and environmental wellbeing. Each category is aggregated from 2-3 subgroups which consist of 2-4 sub-indicators in each.

2.2 Characterising Uncertainty and Understanding Trade-Offs

In order to characterise uncertainties within CIs it first important to understand how CIs are constructed. Although individual CIs differ, Figure 2-1 shows the construction stages for a typical environmental CI. The typical stages of CI construction are:

- **Theoretical framework** is the overarching conception of the CI and choice of subgroups and categories, which act as the key areas of the system that are of interest to be measured. The theoretical framework can impact technical choices such as weighting and normalization.
- **Data selection** involves construction and normalization of variables or sub-indicators as well as analysis and choice of underlying data.
- **Construction** of the CI includes approaches used for aggregation and weighting of sub-indicators, subgroups and categories.
- **Post-development communication** involves dissemination and communication of results.

Many different types and sources of uncertainty exist, emerging from one or more of these stages and requiring different approaches (Figure 2-2). Epistemic uncertainty arises from a lack of knowledge of the dynamics and state of a system and includes uncertainty from limitations of measurement devices, insufficient data, extrapolations and interpolations, and variability over time or space. Linguistic uncertainty is a result of scientific vocabulary being under-specific, ambiguous, vague, context dependent, or exhibiting theoretical indeterminacies (Regan et al. (2002); see Appendix 1, Table A1-1). These uncertainties can be reduced or amplified based on decisions taken during construction (Table A1-2). Explorations of uncertainty in CIs have typically focussed on mathematical rules of construction, primarily related to statistical coherence and precision of the CI, and explored using mathematical techniques such as sensitivity and uncertainty analysis. However, these techniques are still not universally applied. Different construction methods are discussed and summarised in Nardo *et al.*'s (2008) guidance handbook for CI construction. They note the importance of construction decisions, especially an appropriate theoretical framework. Yet they offer little advice to constructors, stating the soundness of the framework and fitness for purpose of the CI is best assessed by the peer community. As such, the theoretical framework for CIs has received less attention in the literature than other sources of uncertainty.

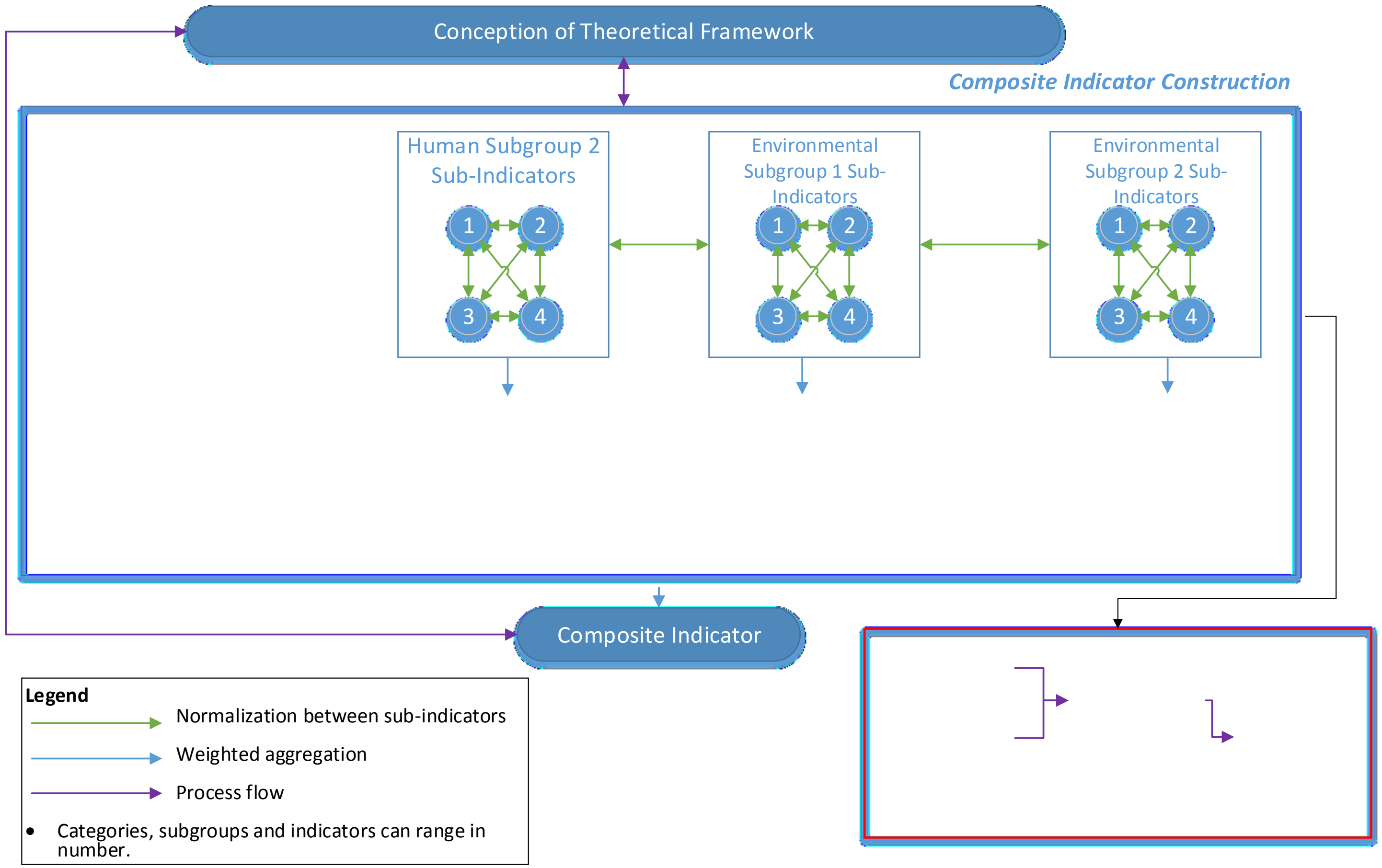

Figure 2-1. Typical construction of a composite indicator. The theoretical framework drives the mathematical construction with some to and fro likely as the index is pieced together. The red box indicates how a typical sub-indicator might be constructed – ideally the desired model would drive the sub-indicator creation, but in reality data availability is often the driving factor. The subgroup boxes show normalization and aggregation of sub-indicators, which in essence create sub-composite indicators.

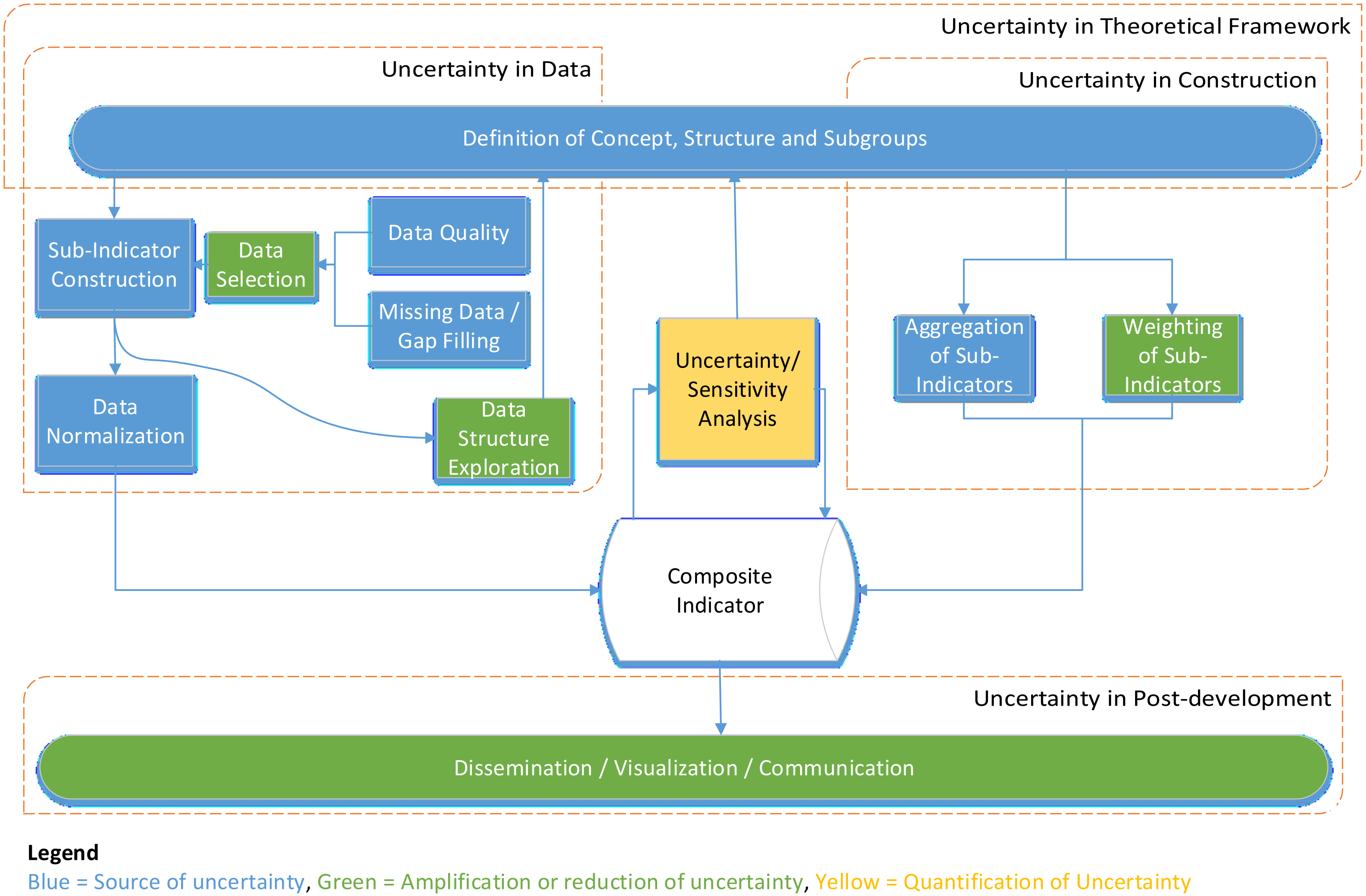

Figure 2-2 Conceptual model of uncertainty flow through a composite indicator. Different stages of construction influence uncertainty in different ways.

2.2.1 Theoretical Framework

The theoretical framework is the starting point for all CIs, and comprises understanding and defining the system to be measured and the contributing categories and subgroups (Nardo *et al.* 2008; Mendola & Volo 2017). For example, the Environmental Performance Index (EPI) has two distinct categories of Ecosystem Vitality and Environmental Health, which contain various relevant subgroups such as ‘Biodiversity and Habitat’ and ‘Air Pollution’ respectively. These subgroups are in this case in essence also CIs, as they are aggregated from several sub-indicators. The final use of the CI is considered here as it affects later decisions on data and construction; for example Buckland *et al.* (2005) suggest criteria for how a biodiversity CI should perform in order to inform choices around construction. Despite numerous CIs existing, little guidance is found in the literature on how to successfully develop a theoretical framework for a CI. General indicator framework advice exists (e.g., using Driving Force-Pressure-State-Impact Response approaches (OECD 2003)), but the distinctive nature of each CI often requires creation of a unique framework that represents the conceptual thinking underlying the indicator. For example, the Ocean Health Index's (OHI) ‘goals’ (categories) were selected based on subject experts reviewing the literature on what the public expects from a healthy ocean. Likewise EPI scanned the literature and policy documents to split their index into two ‘objectives’ and smaller core categories (based on merging the Pressure-State-Response and Driving Force-State-Response frameworks (Hsu *et al.* 2013)). The Sustainable Society Index (SSI) used the Brundtland+ definition of sustainability to pick indicators, which were subsequently grouped into five categories (Van de Kerk & Manuel 2008). Theory of Change (ToC) may prove a useful technique in giving structure to CI design (Weiss 1997). ToC is a useful conceptual tool which works backwards from a desired outcome to consider outputs that will achieve the outcomes and inputs or actions required to deliver these. ToC would help explore the system in question as well as the indicators required to monitor the various outcomes, outputs and inputs.

A clear relationship between what a CI measures and its structure helps provide clarity to the user. However, many environmental concepts (e.g., sustainability) are not well defined, potentially introducing linguistic uncertainty into a CI from the start. This can make it difficult to identify appropriate categories and subgroups. Even widely used terms such as “biodiversity” have associated linguistic uncertainties, which means capturing and monitoring a vague concept can be difficult and highly uncertain (Morar et al. 2015). Lack of clarity causes model uncertainty, as the CI may not actually measure the construct to which it relates.

Once the system has been defined, effort usually focusses on next defining the categories and subgroups which form the structure of the CI. These are outward facing aspects that gain significant attention from constructors as they form the overall communication of the CI, and once selected often act as guidance in determining the indicators that fill them. This approach means that the paths taken to arrive at the final output are often unclear and linkages between subgroups are not fully understood. Relying heavily on a pre-determined theoretical framework rather than developing a conceptual model of the specific processes underlying a given CI runs the risk of arbitrary selection of sub-indicators, which may not be properly representative of the system (Nicholson et al. 2012). Here, ToC could prove particularly useful in providing a transparent and agreed framework to guide indicator selection and reduce the risk of arbitrariness.

2.2.2 Data

As in any quantitative analysis, it is important to understand uncertainties in the data that form the basis of a CI (Figure 2-1 & 2-2). Challenges here include choosing which data best represent components of the theoretical framework, what methods to use to understand uncertain data, and how to deal with the uncertainties.

2.2.2.1 Data quality

Composite indicators almost exclusively use existing data collected by various sources other than the CI creator, over different temporal and spatial scales (Kaufmann et al. 2011). Data are subject to varying levels of uncertainty depending on the credibility of the source, data collection methods, timing of sampling, measurement error, natural variation, and data interpretation. Data uncertainties are rarely adequately quantified (Munda *et al.* 2009), and most uncertainty in CIs is irreducible. At a minimum, a general audit of data quality should therefore be undertaken, with data assessed for relevance, accuracy, timeliness, accessibility, interpretability and coherence before being selected for inclusion (Nardo *et al.* 2008). The EPI and OHI do this by setting rough quality standards for data inclusion, however it is not revealed which data were discarded nor how robust the included data are. Such a process is inherently subjective and users of CIs are not always able to discern where strong or weak data lie. The SSI acknowledges that "*the reliability of data remains a serious concern*", but similarly does not indicate where its strongest or weakest data are found. No discussion of data quality was found for the Climate Change Performance Index (CCPI). Lack of clarity around data that are entered into, or excluded from, a CI might dissuade users or suggest that the CI is a risky basis for policy-making.

Pedigree matrices can be an effective way of assessing unquantifiable uncertainties in data (Van Der Sluijs et al. 2005). This technique involves using qualitative expert judgement to assess parameters through pedigree criteria, which are chosen as the most relevant and applicable criteria to assess parameter strength. Responses are then coded in the pedigree matrix (e.g. between 0 (weak) and 4 (strong)) to reduce arbitrariness and subjectivity. Experts are consulted individually so that consistency across scores indicates a common view of the underpinnings of the parameters, whereas disagreement reflects an ignorance of these underpinnings (Van der Sluijs et al. 2002). This approach thus helps to move consideration of uncertainties beyond those that are quantifiable to the large range of qualitative uncertainties. Assigning data quality scores, as done by the Living Planet Index, is one means of being transparent about data quality issues (Collen et al. 2009). Scores are assigned to data based on their source, methodology and whether a measure of variation was included. Scores can be used either to represent uncertainty, test how overall results differ using different quality data, or adjust weightings for lower quality data. The EPI used a similar method when it was known as the Environmental Sustainability Index, however it was terminated as the method was deemed too subjective and problematic when experts disagreed on assessment criteria for grading (Hsu *et al.* 2013). If data quality is not deemed sufficient then reporting units can be excluded. The EPI did this in their 2012 assessment by removing North Korea, as several anomalous results raised serious questions over data quality.

2.2.2.2 Treating missing data

Data that underpin CIs inevitably contain gaps, requiring decisions about methods used to address these gaps (See Table A1-3). Although modern imputation techniques (such as multiple imputation and maximum likelihood estimation) exist, the use of such techniques may be constrained by time, budget or expertise of the team; these trade-offs need to be reported and justifications given on why certain methods were or were not used. Understanding and displaying where missing data exist is important, as some reporting units (e.g., countries) may be composed of significant amounts of imputed data, which may slip through unnoticed if not transparently logged.

The EPI attempted to collect data for 232 countries but calculations were only performed for 178 due to missing or incomplete data. Conversely, the OHI provides a score for all 221 Exclusive Economic Zones and 15 high seas regions. The OHI fills gaps using a hierarchical decision tree with four different methods: temporal, using data from previous years; alternate datasets used as proxies; spatial, using averages from nearby regions; special rules applicable to particular instances (Halpern et al. 2015c). The SSI has seemingly high data coverage, with less than 10% gaps (Saisana & Philippas 2012), whereas the CCPI offers no discussion of missing data. However, none of the four case study CIs offers an easy insight into which particular data have been imputed, although the OHI has implemented a gap-filling tracking methodology that will be incorporated and presented in global 2016 scores (Frazier et al, in review). CI documentation should be open about which data have been filled, and which reporting regions have been deleted, so the subsequent uncertainty can be properly recognised (Frazier et al. 2016). Regions that contain significant amounts of missing data should be highlighted or removed from the assessment. Analysis of how data gaps affect the overall outcome of the CI is important to guide targeted data collection and inform approaches to gap filling.

2.2.2.3 Data selection and sub-indicator construction

Data selection and sub-indicator construction are intrinsically linked. Ideally, sub-indicators would be selected systematically based on their relevance to what is being measured (Nardo *et al.* 2008; Riedler *et al.* 2015). In reality, the data required to construct an ideal sub-indicator might not be available, be of questionable quality, or have substantial gaps, meaning a trade-off is required. This may involve discarding a preferred indicator in favour of one which is supported by better data, risking introducing severe model error into the CI, or including weaker data in a preferred sub-indicator, meaning it is likely to be less robust.

Typically, individual indicators are chosen using pre-determined selection criteria. Dale and Beyeler (2001) suggest that indicators should be: easily measured; sensitive to stress; respond to stress predictably; anticipatory; predict changes that can be averted; integrative; and have low variability in response to extraneous influences. However, the criteria stated as being important vary widely between indicators. The case study CIs show some consistencies in selection criteria, but also differences (Table 2-1). Lists of selection criteria can enable a more consistent set of indicators, but they do not give much insight into the actual selection process because they do not give information pertaining to why a particular individual indicator or indicator group was chosen and others were discarded, or any relationships between the selected and discarded indicators (Niemeijer & de Groot 2008a). The criteria in Table 2-2 focus mainly on data quality and availability rather than understanding the relevance of indicators to the dynamics of the system, or how indicators might behave as per Dale & Beyler's (2001) list above. Without knowledge of the system's dynamics, it is unclear how the sub-indicators are linked and if they accurately represent the system. Correctly designing sub-indicators is important as they are the basis of a CI; uncertainties here will propagate through the CI, with a 'garbage in-garbage out' logic (Nardo *et al.* 2008).

Table 2-2 Selection criteria for data and sub-indicator construction used in the case study indices

Environmental Performance Index (Hsu *et al.* 2013)	**Ocean Health Index** (Ocean Health Index 2015)	**Sustainable Society Index** (Van de Kerk & Manuel 2008)	**Climate Change Performance Index**
Data must be relevant to what is to be measured	Data must be relevant to what is to be measured	Data must be relevant to what is to be measured	Not located
Indicator provides empirical data on ambient conditions or on-the-ground results for the issue of concern.	Must be able to be scaled to a meaningful reference point	Data must be measurable	

Table 2-2 Selection criteria for data and sub-indicator construction used in the case study indices

Environmental Performance Index (Hsu *et al.* 2013)	**Ocean Health Index** (Ocean Health Index 2015)	**Sustainable Society Index** (Van de Kerk & Manuel 2008)	**Climate Change Performance Index**
Data must have established scientific methodology and based on peer review or institutions charged with data collection	Data must be freely accessible	Data must be from public sources, scientific or institutional	
Must have adequate global and temporal coverage	Must have adequate global and temporal coverage	Must be available for all countries (or all but smallest countries)	
Data represent the best measure available.	Data quality should be considered	Data must be reliable	
Data have been consistently measured across time	Data must be recent and ideally regularly updated	Data must be recent and regularly updated	
		Independent from other indicators with no overlap	

2.2.2.4 Data Normalization

Data that are used in sub-indicators are often in many different formats and therefore must be normalized to the same scale for aggregation (Jacobs *et al.* 2004). This allows comparison of disparate indicators within a single framework. Decisions also have to be made regarding outliers, which can cause problems by becoming unintended benchmarks, skewing data and biasing statistical approaches to weighting. Popular normalization techniques include (Saisana & Saltelli 2011):

- **Ranking** – Simply ranks units in order and therefore does not preserve specific information. Final output is rank only.
- **Standardization (or z-scores)** – converts indicators to a continuous variable with a mean of zero and standard deviation of one. Assumes normality, meaning outliers can have a large effect, which might not be desirable.
- **Min-Max** – normalizes indicators within a given range (e.g. 0-1) by subtracting the minimum value and dividing by the range. Outliers can distort the CI for similar reasons.
- **Distance to target** – Normalizes indicators by dividing the unit's value by a reference target. Can be sensitive to outliers when the best performing unit is used as a target.

The meaning of the results of a CI could be affected by which technique is chosen and should therefore be considered during the theoretical framework stage. Distance to target is a popular method as it allows for the inclusion of political goals, for example the EPI uses the Convention of Biological Diversity's 17% target of terrestrial and inland areas under protection as its critical habitat protection indicator. Normalizing by such a political goal provides a clear benchmark that is relevant and can be easily communicated and frameworks exist for robust selection of quantitative management targets (Samhouri et al. 2012). If using the other more arbitrary methods without testing different techniques, subjective judgement error will mean the outcome of the CI is affected an unknown amount by the choice of normalization.

2.2.3 Construction

This phase determines how normalized sub-indicator scores are aggregated and weighted. Although weighting should be considered during the theoretical framework stage, it is discussed separately here, as it can be an iterative process.

2.2.3.1 Weighting

Weights are often used as measures of perceived importance of the subgroup to the system. For example, the CCPI rates its categories of climate policy, emission trends and emission levels at 20%, 50% and 30% respectively (Burck, J. & Bals 2011). However, lack of knowledge about subgroup importance, or unwillingness to prioritise one area above another, frequently results in equal weight being allocated, which although seen as neutral, is still a weighting decision. The SSI uses equal weights due to a lack of scientific basis for the attribution of weights (Van de Kerk & Manuel 2008). Likewise the OHI uses ten equally weighted 'goals', as the literature does not distinguish which factors are most important for a healthy ocean (Halpern *et al.* 2012), although the OHI explicitly includes a goal weighting term and encourages development and use of weights for localised assessments. The EPI generally sets weights based on the quality of data and relevance of the indicator to the issue it is measuring. Less robust or relevant data are therefore given a lower weighting (Hsu *et al.* 2014). However, despite poor data, the indicator concerned may be key to describing system dynamics; giving it a low weight may therefore reduce the meaningfulness of the CI.

Despite weights often being assigned to different subgroups and/or categories as importance coefficients, variation and correlation in data mean assigned or desired weights might not act as intended. Weighting may have to be an iterative process in order to achieve a desired weighting structure. For example, the EPI performed a sensitivity analysis in 2012 and found that although environmental health and ecosystem vitality had been given equal weights, the greater variation in environmental health scores meant that countries that performed better in environmental health were more likely to perform better in the overall EPI. The weighting was therefore not actually equal and was subsequently adjusted to account for this phenomenon (Hsu *et al.* 2013). However, adjusting weighting to account for variations can be problematic as the assigned weighting is therefore not reflective of the importance, as ecosystem vitality has a larger weighting on face value. Weightings and importance could therefore be misinterpreted by users and therefore should be properly recorded and communicated. Likewise, without such analysis and understanding, weights that are assigned as importance measures may not actually perform as desired in the CI.

2.2.3.2 Aggregation

Two widely used options for aggregation have gained attention in the CI literature; linear and geometric aggregation. Linear aggregation involves a summation of (weighted) sub-indicator scores (usually averaged around a mean), while geometric aggregation involves aggregation by the geometric mean (i.e., using the product of values).

The choice of aggregation method can be a source of model error and subjective judgement uncertainty as it can fundamentally alter how the CI performs. The SSI aggregates through a geometric mean (van de Kerk *et al.* 2014) and the Human Development Index switched from linear to geometric aggregation in 2010 (Klugman 2010), while the OHI, EPI and CCPI all use linear aggregation. Key considerations are:

- ***Compensability*** –Linear aggregation allows complete compensability and geometric partial compensability, which means good performance in one indicator can offset poor performance in another. For example, consider a hypothetical subgroup of a biodiversity index, made up of sub-indicators for protected areas, endangered species and critical habitats (equally weighted). Two countries, A and B, score 9.0, 7.6, 0.4 and 6.8, 6.0, 4.2 respectively. If linear aggregation were used then both countries perform similarly, with a resulting score of roughly 5.6. However, when using the geometric mean, A's overall score is reduced from 5.6 to 3.0, while B's score remains at 5.6. Should no amount of compensability be acceptable and for weights to be truly interpreted as importance coefficients, non-compensatory aggregation methods should be used (Munda 2008; Munda & Nardo 2009; Munda et al. 2009; Cinelli et al. 2014). However, this method has seen limited use and therefore there has been encouragement towards multi-criteria approaches to assess robustness (Munda et al. 2009). Given that many CIs now provide transparency to indicator level, there is less importance placed on the overall score, which can be sensitive to aggregation.

- ***Improvement of scores*** –An improvement in A's critical habitat score from 0.4 to 1.4 in our hypothetical index sees an overall improvement under the geometric mean from 3.0 to 4.6, and 5.6 to 6.0 under the arithmetic mean. The larger jump in the geometric mean could encourage focus on lower performing metrics, which may be beneficial from a policy perspective, but could also dissuade action on higher-performing metrics even if those actions would be beneficial.
- ***Communication*** – linear aggregation is more straightforward for communication and engagement as users can clearly trace scores from the bottom level to the top. It also rewards proportionally to weights, whereas geometric aggregation rewards units with higher scores.

Aggregation to an overall single value is appealing for media traction and communication but may not always be appropriate. For example, the SSI chooses not to aggregate to a single figure based on the strong negative correlation of human wellbeing to environmental wellbeing and thus gives results for three separate composite indicators (the third being economic wellbeing; Saisana & Philippas 2012).

2.2.4 Uncertainty in Post-Development

Communicating a composite indicator can be a complicated undertaking, which can decrease linguistic uncertainty if well done, or increase it by not being transparent and using confusing or vague language. Poor communication may mean a CI is seen as ineffectual or, worse, used improperly. Targeting a technical audience may mean the CI can be critiqued and iteratively improved, but it may not gain the desired public or political traction that a more populist CI would. Likewise, uncertainties can be openly presented or not discussed, but reaffirming complexity within a highly simplified measure is challenging for communication.

A key challenge is whether to focus on the final single numeric output or delve deeper into the CI. By aggregating to a single number, composite indicators can potentially send over-simplistic messages. The same overall score can be achieved in many different ways; one way to overcome this is to give attention to the categories, subgroups and potentially even sub-indicators. All the case study CIs presented here give more information than just ranking countries/overall scores. Detail is given on how scores are achieved, which sub-indicators make up the subgroups and how these have changed over time. The SSI offers downloads of the normalized scores, while the EPI and OHI offer normalized raw data and scores. CCPI offers a qualitative performance review for each country by category but does not discuss sub-indicators or provide data. In all the case studies it is also unclear if or how sub-indicators and subgroups interact, e.g. where an increase in one might cause a decrease in another, although the OHI's inclusion of pressures derived from each category (or 'goal') helps users understand potential trade-offs. Understanding linkages and interactions is potentially critical from a policy point of view, as it is unclear how attempts to alter the status of one particular sub-indicator or subgroup will affect the others. A systems modelling approach could give a starting point for decision makers trying to understand how interactions occur and what their consequences might be.

2.3 Navigating Uncertainty

The range of uncertainties present in CI construction means there is not a single way to treat or represent them all (Table 2-3). Single issue solutions have been addressed in the sections above, but here I lay out potential approaches to improving CI construction that address multiple uncertainties simultaneously.

Table 2-3: Range of uncertainties in composite indicators and how to treat them

Source of Uncertainty	**Issue**	**Reason for Issue**	**Potential solution**
Theoretical Framework	Is theoretical framework representative of the system?	• No systematic process • Subjective • Lack of transparency and repeatability	• Systems Modelling • Systematic expert judgement/stakeholder engagement • Transparency and iterative improvement
Data	Accuracy of data	• Data quality rarely assessed and therefore not really considered	• Data scoring/pedigree matrices • Systematic expert judgement/stakeholder engagement • Uncertainty analysis
	Amount of missing data	• Unclear where data gaps are and number of them. • Gap filling methods are subjective	• Transparency and iterative improvement • Uncertainty/sensitivity analysis • Advanced monte-carlo gap-filling methods
	Is indicator an accurate and desired representation of the system	• Led by data availability, stakeholder or constructor values therefore subjective. • Unclear how indicators relate to system	• Systems modelling

Table 2-3: Range of uncertainties in composite indicators and how to treat them

Source of Uncertainty	**Issue**	**Reason for Issue**	**Potential solution**
	Representation vs quality	• Trade-off between data accuracy and missing data v how well it represents the system • Subjective	• Transparency and iterative improvement • Systematic expert judgement/stakeholder engagement
Data Normalization	Different methods	• Subjective	• Transparency and iterative improvement • Uncertainty/sensitivity analysis
Weighting	Arbitrary weighting	• Unclear how weights were assigned. • "Neutral" weighting still a weighting decision • Subjective	• Systems modelling • Systematic expert judgement/stakeholder engagement
	Implicit weights may be different to assigned weights	• Statistical properties mean assigned weights don't always work as intended	• Correlation analysis • Uncertainty/sensitivity analysis
Aggregation	Different methods	• Subjective	• Transparency and iterative improvement • Uncertainty/sensitivity analysis
Communication	Different interested parties	• How to communicate to public/policy makers/scientists	• Transparency and iterative improvement • Multi-layered approach of engagement/analysis

2.3.1 Systems Modelling

Without a proper understanding of the system and how individual indicators represent its dynamics, CIs risk severe structural uncertainty and improperly informing management decisions. There is a lack of guidance in the literature on how to construct CI theoretical frameworks and as such they tend to follow general approaches such as Driving Force-Pressure-State-Response framework (OECD 2003) or be purpose-built, often informed by literature, experts or stakeholders, focussing on perceived key areas of the system which form the sub-categories (Nardo *et al.* 2008). However, the processes involved in this construction often go undocumented, meaning the resultant structure could be seen as arbitrary and techniques are impossible to replicate. Sub-categories are then populated by relevant sub-indicators. A key feature of any environmental indicator, including CIs, is that it is able to reflect changes in a system. It is therefore crucial that the theoretical framework and selected sub-indicators accurately represent the system. Without first understanding system dynamics, and testing the behaviour of the sub-indicators as the system changes, it is impossible to know if the chosen CI sub-indicators do accurately reflect the system. This means it is not clear whether changes in the sub-indicators are reflecting real system change, or whether critical data gaps exist which could impact on the ability of the CI to track system change.

Systems modelling is an effective approach to representing understanding and thus provides a systematic, transparent and repeatable way to aid sub-indicator selection and theoretical framework development. This approach defines variables or processes which are most important to a system's dynamics, and their interactions, thereby mapping the system and the linkages within it (Niemeijer & de Groot 2008b). Using quantitative modelling approaches to select indicators is well explored in environmental science and is seen as one of the most effective methods of understanding how indicators respond to change. It has been suggested or used as an approach for selecting indicator sets in forest management (Brang et al. 2002; Mäkelä et al. 2012) and commonly used for testing and refining indicators in fisheries (Fulton et al. 2005; Branch et al. 2010). Complex social-ecological models such as ‘Atlantis’ (Fulton et al. 2011a) can provide a basis for understanding systems and picking out key indicators to be included in a CI. Once sub-indicators have been selected, a systems model could be altered based on policy options to test how CIs react to underlying changes in the data; CIs are then able to be validated and act as a decision making aid (Nicholson et al. 2012). However, given CIs often aim to represent highly complex concepts or systems, quantitative models might not be available, or qualitative conceptual and/or expert-based models may be more appropriate.

Qualitative modelling approaches, which map out a system diagram, have proved useful frameworks, for example in indicator selection for sustainable tourism (Margoluis *et al.* 2009;), fisheries (Vugteveen *et al.* 2015), land use change (Benini et al. 2010; Van Oudenhoven et al. 2012), land degradation (Gisladottir & Stocking 2005; Agyemang et al. 2007), urbanization (Jago-on et al. 2009) and water management (Chung & Lee 2009). Conceptual modelling like this allows stakeholders to come together and help develop the model, which enables sharing and inclusion of perspectives and values. A conceptual systems model can be used to select indicators by first defining clearly a concrete question to be answered. In the case of CIs, this is likely to be broad (i.e. measuring sustainability of countries) but could also be very specific (i.e. the effect of increasing protected areas on biodiversity). Broader questions will naturally require indicators that span multiple issue areas and will not be as fine-tuned as those that are selected to help track more specific questions; indeed, specific questions may require more detailed construction of the systems model. In order to select indicators, nodes must first be identified: Root nodes which have many outgoing arcs typically provide information on sources of issues; central nodes with many incoming and outgoing arcs are usually important for the most general of indicators as they provide information on many issues; end of root nodes with many incoming arcs allow the gauging of multiple issues at once. These nodes can then provide guidance on the types of indicators that are required to answer the question. Niemeijer & de Groot (2008a) provide a detailed example using this approach to select indicators for ecological impact of nitrogen fertilization on surface waters. Such an approach could then allow CIs as a means to track specific issues, by using the systems model to pull out sub-indicators relevant to a specific issue. A systems model means indicators are not selected arbitrarily, have a wider and understandable function and can be interpreted effectively, minimising the risk of incorrect analysis.

A system model could also inform weightings by pinpointing those system areas that are more central or peripheral, meaning final weights could be selected by an informed systems approach, thereby constraining uncertainty. Selected sub-indicators found to be suboptimal due to data issues could be highlighted, aiding in representing uncertainty as well as showing key data gaps that should be filled in order to have a more robust understanding of the system. Starting with a systems approach could help balance the current top-down framework creation and reduce arbitrary indicator selection.

Of course, in a CI, subgroups are often based on stakeholder values or perceptions, and this is recognised as an important facet of CIs. While stakeholder engagement is always important, it is particularly critical where highly complex models include large amounts of uncertainty or system dynamics are unclear. I therefore encourage an iterative back-and-forth approach between the modellers who can point to expected important variables based on their system models and stakeholders who can likewise do the same based on societal values.

2.3.2 Systematic expert judgement and stakeholder engagement

Expert-led or stakeholder-participation approaches and are often used for CI construction, but with little information supplied on how or why decisions were made. This lack of transparency means the results of engagement are often unknown and methods unrepeatable. Structured elicitation of knowledge from experts is well explored and can be a powerful tool if used correctly. Key lessons include; elicit knowledge from groups rather than individuals, carefully choose members, strive for group heterogeneity, calibrate and weight experts, train experts, and give feedback (Burgman et al. 2011a, 2011b; Sutherland & Burgman 2015). Using these techniques opens up many options throughout constructing a CI; such as providing judgement on theoretical framework and sub-indicator selection, estimating data accuracy by providing bounds or data scoring, providing guidance on weighting and help with communication and analysis of results. McBride *et al.* (2012) demonstrate how such techniques were applied to the IUCN Red Listing of Australian birds in order to reduce bias and error amongst experts.

Furthermore, their elicitation was carried out online, showing that lengthy and costly workshops do not always have to be undertaken.

Stakeholder engagement is important for ensuring that the CI is useful for the intended audiences, and should be done at an early stage and throughout CI development and implementation. The more diverse the perspectives involved in developing the conceptual framework and exploring sub-indicator selection, the more likely it is to represent meaningful reality for the end users (Burgman et al. 2011b; Fulton et al. 2011b). This engagement can be more difficult when large numbers of different types of stakeholder are involved. Even at the global scale, representatives of particular groups can be consulted. Stakeholder input can be particularly vital, however, when CIs are used at more regional or local scales. Systematic, recordable techniques are useful in order to document engagement outcomes. Halpern *et al.* (2013) used such methods (based on random utility theory and analytical deliberation) to elicit stakeholder preferences for indicator weighting in the OHI in a regional assessment of the California Current. However, such a task was considered by Halpern *et al* (2013) to be unworkable on a global level as the range of preferences would be so vast. Indicators tend to be less successfully utilised when they are purely scientific; involving stakeholders and leaving room for negotiation in CI construction can be highly beneficial in the messy situations CIs tend to be needed for (Turnhout et al. 2007).

2.3.3 Statistical coherence and robustness

The combined use of uncertainty and sensitivity analysis for CIs is well explored within the literature but is still not universally applied (Saisana *et al.* 2005; Munda *et al.* 2009; Paruolo *et al.* 2013). Uncertainty analysis focuses on how uncertainty in inputs, such as poor data or subjective construction choices about aspects like the weighting scheme, propagates through the CI to affect outputs. Results are usually represented by uncertainty bounds around output values. Sensitivity analysis looks at how each individual source of uncertainty contributes to this variation, and has been used to investigate the robustness of several CIs, including the

SSI and EPI. The latest report on the EPI (Athanasoglou et al. 2014) found that three of the nine issue areas did not contribute significantly to EPI ranking, suggesting that changes to these indicators should be made. Aggregation function choice was found to account for 94% of sample variance, whilst choice of weighting for the objectives only accounted for 4%. This suggests that further discussions surrounding methodological choice should focus on aggregation method rather than weighting. Ninety percent of SSI countries shifted less than ±1 position with respect to the simulated median, suggesting that the 2012 SSI is not unduly driven by methodological assumptions (Saisana & Philippas 2012). The OHI and CCPI have yet to undertake uncertainty/sensitivity analyses, although they are planned for the OHI.

The combined use of uncertainty and sensitivity analysis provides an evaluation of confidence in the mathematical properties of the CI, assessing uncertainties associated with the construction process. This can help with many of the methodological decisions I highlight. However, uncertainty and sensitivity analyses only deal with a limited part of the uncertainty surrounding CIs. For example, they can show how rankings change based on the methodological choices made, or if any bias is present. They cannot, however, detect whether indicators are measuring what they intend to, or whether the CI represents overall system dynamics.

2.3.4 Communication and transparency

A key benefit of CIs is their ability to communicate issues clearly, to a wide audience, by aggregating sub-indicators to a single figure. However, many CIs act as more than a communication tool and are used for tracking trends and decision making. This simplistic output of a single number or score then becomes problematic, as some parties will not believe in complete aggregation as an approach to summarising complex and interacting systems. The deeper a CI can be explored, the more useful it will be for technical audiences. Therefore, complete transparency to sub-indicator level, including documentation of the issues covered here is desirable, which is not usually the case (Freudenberg 2003; Munda 2005; Böhringer &

Jochem 2007; Singh et al. 2009). Users such as scientists and management authorities may require different or more detailed information, particularly related to how sub-indicators were selected, how they interact, where data gaps are found and why methodological decisions were made. This information is necessary for management, can ensure CIs are quality checked appropriately through peer review, and supports their iterative development. Transparency should involve acknowledging the methodological decisions made during construction and the rationale for employing certain methods over others.

All CIs in Table 2-1 provide detailed documentation on their methods and results in an accessible format through their website. Currently, however, none of them discuss uncertainty beyond the uncertainty and sensitivity analysis performed for the EPI and SSI. This limits constructive criticism and improvement. Given the variation in sources of uncertainty and methods of treating them, communicating these effectively becomes problematic. Those who solely use CIs for their most simplistic numeric output are unlikely to be interested in the technical uncertainties. Attempting to communicate these uncertainties might dilute the effectiveness of CIs themselves. Therefore, although uncertainty considerations are critical, I believe they should be reserved for more technical audiences.

2.4 Future Considerations

Environmental CIs are increasingly produced, but have often been criticised for their lack of acknowledgement and treatment of uncertainty (Böhringer & Jochem 2007; Jørgensen et al. 2013; Giampietro & Saltelli 2014). I have provided a comprehensive assessment of the sources of uncertainties and methods to treat, represent or reduce them. Articulation and treatment of uncertainty within CIs is underdeveloped; how it is accounted for will depend on individual CIs' aims and audiences. Transparency and acceptance of uncertainty may be sufficient in some CIs, but others may require reworking of sub-indicators and construction methods. I hope this framework can act as a basis for considering uncertainty at each stage and recording not only why certain decisions were taken, but why others were not. Choices are unavoidable during CI construction but should always be acknowledged.

The focus to date on mathematical techniques for dealing with uncertainties in CIs has meant the role of the theoretical framework and sub-indicator selection has received less attention. Approaches such as systems modelling are fundamental for proper selection and grouping of indicators and their interactions, if a CI is properly to represent reality. This would also increase their usefulness in a policy setting, by testing policy scenarios and selecting specific sub-indicators to help answer more explicit questions. Importantly, development of a CI should be an iterative process. Expressing the location and importance of different types of uncertainty can then be a catalyst for new data collection or conceptual development that is targeted at reducing the most influential uncertainties. Communicating these uncertainties may need to be done separately from the public-facing communication of the main CI, so as to not dilute its impact. However, there is much scope for novel, simple and clear techniques for communicating uncertainties effectively to technical audiences who require such information. This will build trust in CIs and thereby enhance their ability to support decision-making.

3 A PAN-ARCTIC ASSESSMENT OF THE STATUS OF MARINE SOCIAL-ECOLOGICAL SYSTEMS

3.1 Introduction

Arctic ecosystems are experiencing profound physical, ecological and social changes, driven largely by a warming climate and increasing economic development (Hovelsrud et al., 2011; Wassmann et al., 2011). There is a need to establish a baseline of the biophysical and socio-economic dimensions of ocean ecosystems across the Arctic, which can then be used to assess the consequences of future change (IOC/UNESCO 2010). Such baseline assessments are necessary to support strategic and evidence-based decisions for conservation and economic investment through Ecosystem-Based Management (Elliott 2014). Tools such as the Ocean Health Index (Halpern et al. 2012) can help provide a framework for collating and analysing a wide breadth of baseline data to facilitate management (Borja et al. 2016).

Each Arctic state above the Arctic Circle (Russia, Canada, USA, Norway, Denmark [Greenland]) responds to and manages its Arctic areas through its own national governance system. However, many issues are transboundary in nature, requiring co-management and collaboration (Van Pelt et al. 2017). For example, Arctic ecosystems support globally significant fisheries, with many species already undergoing range shifts and changes in abundance, growth and phenology (Wassmann et al. 2011; Pinsky et al. 2018). Also, shipping through the Northern Sea Route is increasing annually (Northern Sea Route Information Office 2013), with future projections signalling exponential increases (Smith & Stephenson 2013). Yet while there is some international cooperation through bodies such as the Arctic Council, there have been few legally binding commitments across nations to collectively and systematically manage the challenges facing the Arctic marine areas. Such examples are limited to international agreements on oil spill preparedness, search and rescue at sea, and the Oslo Declaration preventing fishing in the currently ice-covered central Arctic Ocean (Baker & Yeager 2015; Molenaar 2015). A review by Protection of the Arctic Marine Environment (2013), an Arctic Council working group, found that there was a need for further coordination across institutions (e.g., monitoring conducted on a polar-wide basis using consistent methods, with central data storage) and further cooperation and knowledge-sharing between Arctic countries and institutions. The current disconnect in monitoring and policy at the pan-Arctic scale 'limits the ability to efficiently make effective management decisions' (CAFF 2014). Ultimately it recognised that there is a need to amend existing instruments or develop new ones to strengthen governance for the conservation and sustainable use of the Arctic marine environment.

The Ocean Health Index (OHI) is a tailorable marine assessment framework to comprehensively and quantitatively evaluate ocean health (Halpern et al. 2012; ohi-science.org). It is increasingly being used to help guide thinking around marine management, particularly in data-limited areas, by providing a structure to analyse data availability (Lowndes et al. 2015). I performed an OHI assessment for the Arctic to bring together disparate data and establish an initial baseline of social-ecological conditions, with a focus on highlighting areas of potential concern (both geographically and by goal), exposing data uncertainties, and highlighting potential interactions between marine management goals and short and long-term outcomes. I discuss the results in the context of future management and decision-making in the pan-Arctic region. Like the OHI (Lowndes et al. 2017), the Arctic OHI (AOHI) is a flexible framework with accompanying open software, and can be iteratively improved over time as better data becomes available or stakeholder values are more comprehensively included.

3.2 Methods

3.2.1 Pan-Arctic Region

Many different definitions of the Arctic exist, with boundaries defined by physical delineations (e.g., climate), latitude, extent of continuous permafrost or sea ice, treeline, or geopolitical borders (Maher 2007). Indeed, spatial delineations of the Arctic even differ between Arctic Council Working Groups (Koivurova 2010). I used Exclusive Economic Zone (EEZ) boundaries that fall above the

Arctic Circle as the primary filter for developing an Arctic OHI (AOHI) that can feed into national and international monitoring and policy efforts across the Arctic, with the exception of Southern Greenland which was included because data reported for Greenland often included this area (Figure 3-1; Table A2-1). I excluded the high seas regions because I chose to focus on comparing national Arctic EEZs for management potential. I call this case study region the pan-Arctic area.

I further subdivided the Norwegian, Greenland and Canadian EEZs based on defined management areas and scales of data reporting. I could not subdivide Russia's Arctic region, despite it being the largest of all countries, due to Russia's marine governance structure, which is managed centrally and thus limited data was available at sub-national scale for many goals.

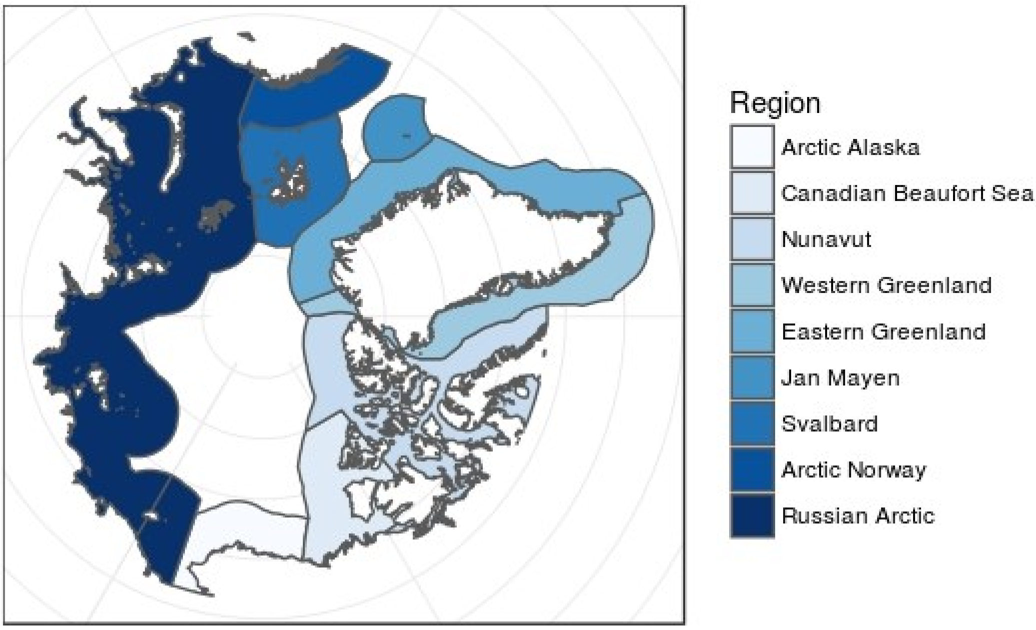

Figure 3-1: Pan-Arctic Region

3.2.2 Index Calculations

The OHI framework and methods are detailed extensively in the literature and public domains (Halpern et al., 2012, 2015, 2017; ohi-science.org); here I present a brief overview and focus on changes and new approaches instituted for the AOHI. The OHI is based on assessing the status of an area against a set of goals, which represent key values and benefits people hold and want from a healthy ocean (Table 3-1). A healthy ocean is defined as one that sustainably delivers a range of benefits to people now and in the future (Halpern et al. 2012). A key design element of the OHI is that it can be adapted to fit different scales and incorporate different values and benefits into its goals, depending on the location and aim of the assessment (Lowndes et al. 2015); for example, OHI assessors determine the number of goals to be evaluated depending on the region of interest and the aspects of ocean health relevant to that region. Numerous OHI assessments have been completed all over the world (Halpern et al. 2013b; Elfes et al. 2014; Selig et al. 2015; Daigle et al. 2017; Longo et al. 2017), with many more in progress. Relative to the global assessment, localised assessments are able to take advantage of higher-resolution data, more locally relevant reference points, and goals adapted to local values (Daigle et al. 2017). Where localised data are unavailable, regional assessments can use existing country-level data from the global OHI.

The OHI is calculated by combining individual indicators via a structured framework designed to measure progress toward optimal sustainable delivery of each of the goals (four of which are further subdivided into sub-goals). For the AOHI, I assess 9 of 10 goals in the global OHI, with Carbon Storage not assessed due to lack of data for calculating a meaningful indicator (Table 3-1). Each goal and sub-goal is measured on a scale of 0-100, with 100 being the highest possible score. Each goal score, G_i, is calculated as the average of the current status, x_i, and the likely future status, $\hat{x}_{i,F}$:

$$G_i = \frac{x_i + \hat{x}_{i,F}}{2}$$

(Eq. 1)

Current status measures the most recent year's performance relative to a reference point of the highest sustainable performance for that goal. Likely future status captures the near-term (five years from current status) future performance for the goal based on recent trend in status, (T_i, calculated as the slope in the change of the status score of the previous five years), pressures that can threaten the delivery of each goal p_i), and resilience factors (r_i) which can mitigate these pressures.

$$\hat{x}_{i,F} = [1 + \beta T_i + (1 - \beta)(r_i - p_i)]x_i$$

(Eq. 2)

Following Halpern et al. (2012), β represents a weighting factor of 0.67, giving trend twice the importance compared to pressure and resilience terms, reflecting the better indication of near-term trajectory that trend provides. Resilience typically measures policies or international conventions to which regions are or are not party, but also includes ecological and socio-economic resilience. For climate change-related pressures, resilience layers were set as zero for the AOHI, as no resilience measures adequately offset the pressures of climate change in the Arctic (Bennett et al. 2015). See Table A2-2 and A2-3 for pressure and resilience matrices.

I calculated the overall regional AOHI scores as an equally weighted average of goal scores because, given the heterogeneity of people and environments across the Arctic, determining weightings would be a substantial undertaking and outside the scope of this study. Furthermore, I focus on results comparing goals and regions to avoid focus on higher aggregation and weightings. Weightings could be altered in the future based on stakeholder consultations (Halpern et al. 2014; Daigle et al. 2017). Finally, the overall AOHI score was derived using an area-weighted mean of scores for each region within the assessed area. Below I describe the data and methods used to calculate AOHI scores. All original data, scripts used for processing, final data layers and goal models are open access and freely available online at https://github.com/OHI-Science/arc. The AOHI was calculated using the R package 'ohicore' (Ocean Health Index 2016).

Table 3-1: Goal by goal overview of the Arctic OHI. Details of any differences from the global OHI are discussed in the text.

Goal	Sub-goal	Reference Point	Model different from Global OHI?	Data	Regions included*
Food Provision	Fisheries	All stocks fished at levels that afford maximum sustainable yield.	N	Spatial fisheries data from the Sea Around Us Project (Pauly & Zeller 2015)	All regions
	Mariculture	95th percentile of best performing region	N	Norway production data from Arctic counties. Russian Arctic production data from Food and Agriculture Organisation (FAO 2011-2017).	Arctic Norway, Russian Arctic
Clean Waters		Zero nutrient and chemical pollution, pathogens and marine debris	N	Spatial data for each pollution type from global assessment (Halpern et al. 2015)	All regions
Coastal Economies and Livelihoods	Coastal Economies	Moving window of revenue values over previous 5 years	N	Marine sector revenue data for each region, limiting to Arctic areas where possible. Revenue data purchasing power parity (PPP) adjusted for comparison across regions	Arctic Alaska, Canadian Beaufort Sea, Nunavut, Russian Arctic, Arctic Norway, Svalbard, W. Greenland, E. Greenland.
	Coastal Livelihoods	Moving window of number of jobs and average wages over previous five years.	N	Marine sector employment and wage data found for each region, limiting to Arctic areas where possible. Wage data PPP-adjusted for comparison across regions	Arctic Alaska, Canadian Beaufort Sea, Nunavut, Russian Arctic, Arctic Norway, Svalbard, W. Greenland, E. Greenland.
Sense of Place	Iconic Species	All iconic species at Least Concern on IUCN Red List	N	Pan-Arctic iconic species; those included on the WWF Iconic Species list and the Arctic Biodiversity Trends Indicator Species.	All regions

Table 3-1: Goal by goal overview of the Arctic OHI. Details of any differences from the global OHI are discussed in the text.

Goal	Sub-goal	Reference Point	Model different from Global OHI?	Data	Regions included*
				Local iconic species selected for each region through literature review.	
	Protected Places	30% of coastal areas (within 3nm and 1km inland) under protection (as defined by World Data Base for Protected Areas (WDPA)	Goal definition changed	WDPA (IUCN & UNEP-WCMC 2017)	All regions
Coastal Protection		Average extent of shoreline sea ice 1979-2000	N	Spatial shoreline sea ice data from NSIDC. (Cavalieri et al. 2015)	Arctic Alaska, Canadian Beaufort Sea, Nunavut, Russian Arctic, Arctic Norway, Svalbard, W. Greenland, E. Greenland.
Marine Mammal Harvest		Sustainable harvest of marine mammals: Catch/Catch Limit = 1	Y, replaces global Natural Products goal	Marine mammal harvest data and corresponding quota or potential biological removal for each region.	Arctic Alaska, Nunavut, Russian Arctic, Norway, Jan Mayen, W. Greenland, E. Greenland.
Biodiversity	Habitats	Sea ice: No loss of sea ice habitat compared to average extent 1979-2000 Soft bottom: Inverse relationship to 95th percentile of highest global trawl density	N	Sea ice: Spatial sea ice data from NSIDC (Cavalieri et al. 2015) Soft bottom: Trawling density maps (Halpern et al. 2015)	All regions
	Species	All assessed species at Least Concern on IUCN Red List	N	IUCN Red List spatial data (IUCN 2017)	All regions

Table 3-1: Goal by goal overview of the Arctic OHI. Details of any differences from the global OHI are discussed in the text.

Goal	Sub-goal	Reference Point	Model different from Global OHI?	Data	Regions included*
Artisanal Needs		Sea Ice: Average overall sea ice extent 1979-2000 Marine Mammals: Artisanally targeted marine mammals all Least Concern on International Union for Conservation of Nature (IUCN) Red List Fisheries: Artisanally targeted species have sustainable stocks ($B/B_{MSY} = 1$).	Y, goal altered	Sea Ice: Spatial sea ice data from National Snow and Ice Data Centre (NSIDC) (Cavalieri et al. 2015) Marine Mammals: IUCN Red List. (IUCN 2017) Fisheries: Artisanal catch data from Sea Around Us Project. (Pauly & Zeller 2015)	Arctic Alaska, Canadian Beaufort Sea, Nunavut, W. Greenland, E. Greenland, Arctic Norway, Russian Arctic
Tourism		90^{th} percentile of region best performer	N	Tourism employment data	Arctic Alaska, Canadian Beaufort Sea, Nunavut, Russian Arctic, Arctic Norway, Svalbard, W. Greenland, E. Greenland.
*Regions are excluded from certain goals due to lack of relevance for that region (i.e. Jan Mayen is uninhabited and so many human-oriented goals are not included)					

3.2.3 Arctic Ocean Health Index Goal Calculations

Given the scale and heterogeneity of the region and to facilitate comparisons across these scales, most goal models and reference points were not changed from methods used in global assessments (Halpern et al. 2017), except for two new goals, Marine Mammal Harvest and Artisanal Needs, which were adapted Natural Products and Artisanal Opportunities goals to better suit the Arctic region (Table 3-1). An overview and key details on all goals are provided below (with expanded detail in Appendix 2); greater detail is provided for the two goals that were adapted for this assessment. Data sources are listed in Table 3-1 and a full list of data layers can be found in Table A2-4.

The AOHI focusses on the Arctic region, including partial coastlines of several nations. This scale results in added complexity because many data sources are reported at national-level resolution, and so cannot be directly used. For example, data for the entire USA are not representative of the North Arctic Alaskan coast, as they represent all USA regions. Furthermore, the large study area and heterogeneity of the region made obtaining data challenging; when local data were not available or not comparable across all Arctic regions, I often used global spatial data refined to the Arctic region (detailed below). In tailoring the assessment to the Arctic, I was able to replace or adapt 74% of the data layers (n=81) to be specific to the Arctic. The unchanged 26% mainly consisted of resilience scores for national-level factors, such as whether each country was a signatory to the Convention on International Trade of Endangered Species (CITES).

Due to difficulties and gaps in monitoring the Arctic, pan-Arctic datasets are likely to contain many uncertainties or errors which could affect the results of the AOHI. It is beyond the scope of this work to fully assess and account for these possible sources of uncertainty (Burgass et al. 2017). As an example to help illustrate and understand how uncertainty might affect results, I recalculated the Fisheries sub-goal of Food Provision using a different source dataset for fisheries catch.

3.2.3.1 Food Provision

This goal intends to capture whether seafood provisioning potential is sustainably maximised in each region, through both wild harvest (Fisheries sub-goal) and cultivation (Mariculture sub-goal). The two sub-goals are combined to give an overall score for Food Provision, weighted by their tonnage contributions.

Fisheries

The Fisheries sub-goal is based on the amount of wild-caught seafood that is sustainably caught within the study area. I used data for this goal taken from the Sea Around Us Project (www.seaaroundus.org), which reconstructs catch data and spatially distributes catch across the world at half-degree resolution (Watson et al. 2004; Pauly & Zeller 2016). In line with (Halpern et al. 2015c), I used catch data to calculate B/B_{MSY}[1] as a measure of stock status when stock assessments were not available, and penalized scores when taxa were not reported at species level to highlight a potential lack of adequate species-level management. Finally, I calculated overall status as the mean of the stock status scores, weighted by the average overall catch in that area, across the time series (see Appendix 2). I also used an alternative fisheries dataset, from Watson (2017) to test the sensitivity of results to the data used. These data are spatially disaggregated catch data at 0.5 degree cells, similarly presented to the Seas Around Us dataset. As such I processed the data in a similar manner and ran this through the AOHI to see how scores might change.

Mariculture

The Mariculture sub-goal assesses the sustainability and production of ocean-farmed seafood. Mariculture currently only occurs in Norway and north-west Russia, which were the only two regions to include this sub-goal. I estimated sustainability of production based on Trujillo (2008), as has been done in other OHI studies (see Appendix 2). The goal model

[1] For a particular fish stock, the ratio of observed biomass (B) to the biomass that would provide maximum sustainable yield (B_{MSY}). When $B/B_{MSY} = 1$, then biomass equals B_{MSY}. If B/B_{MSY} falls below 1, biomass is too low to provide maximum sustainable yield.

calculates status as the mariculture yield multiplied by sustainability coefficient and normalised by coastal population, which is necessary for undertaking aquaculture. Arctic aquaculture is limited and a global reference point would not be appropriate, so a regional reference was set as the 95th percentile of the top performing region, Norway. While this currently only includes two regions, the expansion of aquaculture in to new regions in the future means the same reference can be used for repeat assessments.

3.2.3.2 Clean Waters

The objective for the Clean Waters goal is to maintain the ocean free of contamination, pathogens and anthropogenic nutrient enrichment, for both recreation and environmental health. This goal used four types of pollution data: trash (marine plastics), chemical (runoff, shipping and ports), pathogen (sewage waste) and nutrients (land-based inputs). I refined each of these global data layers (Halpern et al., 2015) at the 1km^2 raster level to only include areas within the AOHI and scaled each raw pollution data layer from 0 to 1, with 1 indicating the highest level of global pollution (Halpern et al. 2015a). I calculated goal status for each region by determining mean rescaled score for each pollution type, subtracting the mean rescaled pollution scores from 1, and combining the four scores using a geometric mean.

3.2.3.3 Coastal Livelihoods and Economies

This goal tracks the number and quality of jobs and the amount of revenue produced from marine-related industries and sectors through two sub-goals, Livelihoods and Economies. A score of 100 reflects productive coastal economies that avoid the loss of ocean-dependent livelihoods while maximizing livelihood quality.

Livelihoods

This sub-goal describes livelihood quantity and quality for people living on the coast. The livelihoods subgoal includes two equally weighted sub-components; the number of jobs, which is a proxy for livelihood quantity, and the per capita average annual wages, which is a proxy for job quality. I obtained job and wage data for marine sectors from within each region at as fine a scale as possible (Table A2-7 and A2-8), and then aggregated it by region (e.g. Arctic Alaska data aggregated from Northwest Arctic Borough and North Slope Borough statistics). Wages were then adjusted using Purchasing Power Parity (PPP) which enables direct comparison between nations and across years. As per Halpern et al. (2017), the reference point for jobs is a temporal comparison using a five-year moving reference value, in which a score of 100 indicates that the number of marine jobs in a given area has not declined relative to five years previously. Similarly, for wages, a score of 100 means the adjusted wage has not declined relative to the highest average annual wage observed across all reporting units five years previously.

Economies

The Economies sub-goal captures the economic value associated with marine industries based on reported revenue from marine sectors. I obtained revenue data for marine sectors across the Arctic (see Table A2-9). Values were adjusted by PPP, and as for livelihoods, the reference value was a moving target temporal comparison. A score of 100 indicates that revenue has not decreased compared to its value five years previous.

3.2.3.4 Sense of Place

The Sense of Place goal aims to capture the desire to preserve areas and species that contribute to peoples' connection to the oceans. This connection might arise from socio-cultural values which local communities have for traditions tied to the existence of these places or species, or from the existence of species or locations that are iconic to a wider public, though they may never be experienced directly.

Iconic Species

Iconic marine species fell into three categories: 1) Those that are considered globally iconic were taken from the WWF Global Priority and Flagship species list and are selected for all regions in which they occur, or 2) Those that were considered regionally iconic were taken from CAFF (2010) who identified species that were of wide relevance to indigenous or local people across the Arctic and 3) Those that are locally iconic for cultural or social reasons, which may differ between regions. Locally-iconic species were selected based on a review of grey literature to help synthesise and identify which species were culturally important for each region. Determining what is iconic is ultimately subjective, but methods such as those used by Daigle et al. (2017), who defined iconic species as those that appear on Canadian coins in their Canadian OHI assessment, or Roll et al. (2016) who investigated cultural importance of reptiles through internet interest, show that more systematic methods are possible. Iconic species selected for each region can be found in Table A2-9. The average conservation status of these iconic species (from the International Union for Conservation of Nature (IUCN) Red List) was converted to a numerical score (see Table A2-13), with the reference point equal to having all these species listed as Least Concern. Given the plethora of species that could potentially be considered iconic across the Arctic, a more systematic and locally-driven selection process would be beneficial in the future.

Protected Places

I altered the name of this goal slightly to focus on 'Protected Places', rather than 'Lasting Special Places' found in other OHI assessments, although it is assessed in the same way. This name change was due to the lack of information related to what areas might be considered special in the Arctic, and the fact that many of these areas might be unsuitable for protection given their cultural importance as fishing and hunting grounds. I scored this goal by comparing the amount of protected area within 3 nautical miles offshore and 1 km inland, as defined by the World Database on Protected Areas (WDPA), compared to a reference point of 30% of the total area protected.

3.2.3.5 Coastal Protection

This goal assesses the amount of protection provided by marine and coastal habitats against erosion to coastal areas. In the global assessment, saltmarsh, coral, mangrove and sea ice habitats are all assessed, while in the Arctic the main habitat offering coastal protection benefits is sea ice. In the AOHI I assessed the current amount of shoreline sea ice (averaged over the previous 3 years to help reduce the impact of natural variation) compared to the reference condition (average sea ice extent between 1979-2000).

3.2.3.6 Marine Mammal Harvest

This goal is analogous to the Natural Products goal that is assessed globally, in that it aims to measure the sustainable harvest of non-food marine resources – in this case the harvesting of marine mammals for furs, ivory and other resources, which is an important activity commercially and culturally across the Arctic (Hovelsrud et al. 2008). While marine mammals are also eaten, this is not a sole reason for their exploitation and thus I do not count marine mammal harvests under the Food Provision goal. Accurate and repeated measures of sub-population sizes are not available for many Arctic marine mammals, which are actively hunted, making construction of a meaningful indicator challenging (Laidre et al. 2015). I therefore only considered species for which either a quota (assumed to be sustainable) or potential biological removal (PBR) rate was available (Table A2-12); these tended to be pinniped species which haul out of the water, making population estimates easier. Several whale species are hunted by Arctic communities under aboriginal quotas issued by the International Whaling Commission (IWC). These are issued over a five-year period, rather than annually, and the current quotas are operational until 2018, at which point they can be properly assessed for under- and overharvest. I do not include these whale species in this iteration of the AOHI, but assessing a wider array of species than just pinnipeds would provide a more comprehensive understanding of the status of marine mammal harvest across the Arctic. I calculated the goal status as the ratio of current harvest compared to the current reference point (quota or PBR), similar to the natural products goal for the Southern Ocean OHI assessment (Longo et al. 2017).

A Catch per Catch Limit score (*C/CL*) was initially calculated to determine landings relative to the quota or PBR (catch limit) for each region and year:

$$C/CL = \frac{catch}{catch_{limit}}$$

(Eq. 3)

These values were then converted to a stock status score (*S'*), which ranges from 0 to 1 (Figure 3-2), and penalized for over- and under-harvesting (although over-harvesting is more harshly penalized). The lowest value that can be obtained when the catch is lower than the catch limit (i.e. under-harvest) is 0.25 as under-harvesting can be beneficial to rebuild populations. A buffer range of 0.9 to 1.1 was established around a C/CL score of 1.0 to account for uncertainty and fully reward regions aiming to meet quotas. If a region contains more than a single species of hunted marine mammal, then scores were averaged across species.

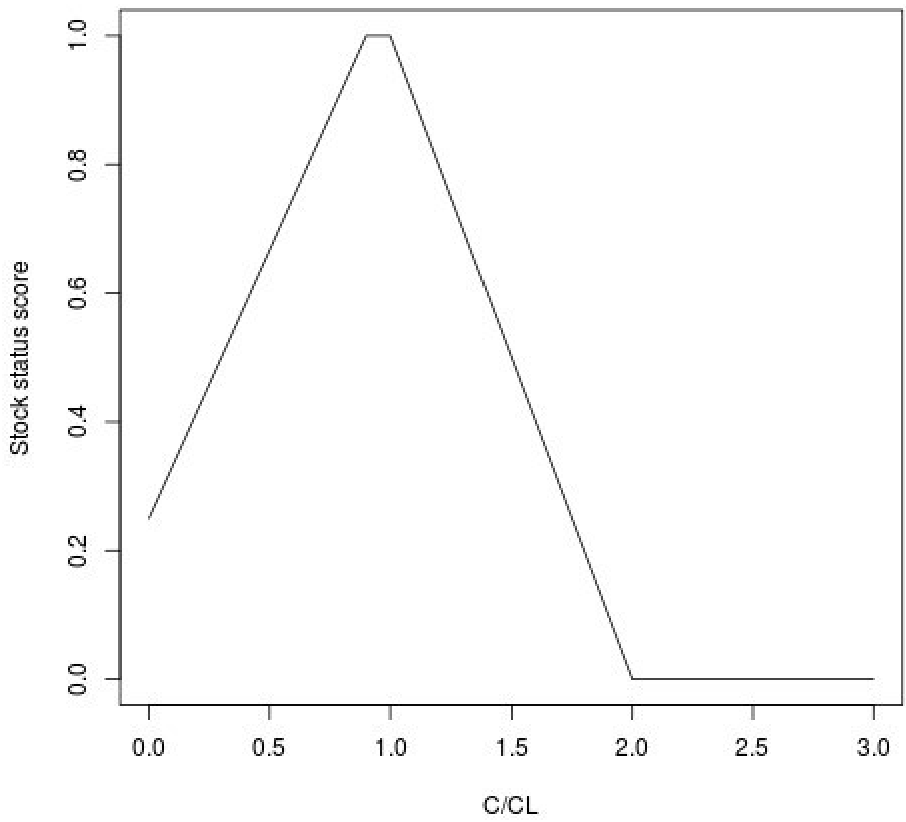

Figure 3-2: How stock status score was generated from the Catch per Catch Limit score (Equation 3) for Marine Mammal Harvest

3.2.3.7 Biodiversity

The Biodiversity goal captures the preservation of biodiversity for its aesthetic, existence, and supporting service values into the future. Biodiversity is measured through two proxy sub-goals, habitats and species. Monitoring biodiversity on a pan-Arctic scale until recently has been disjointed and non-standardised, meaning I relied heavily on global data. With the launch of the Circumpolar Biodiversity Monitoring Plan, this is a goal which hopefully can be improved in the future (M.J. Gill, et al. 2011) .

Habitats

Soft-bottom habitat and sea ice were the only habitats for which data exist across the entire assessment region, signalling a requirement for greater monitoring of Arctic marine habitats. Habitat condition for soft bottom subtidal habitat was estimated using a proxy based on the intensity of trawl fishing relative to soft bottom subtidal habitat area (see Appendix 2). Sea ice condition was estimated by comparing the average extent of all current sea ice edge (averaged over the previous 3 years), compared to the reference point of the 1979-2000 average extent.

Species

Species status data come from the IUCN Red List, which assesses entire taxonomic groups in categories of threatened status. Thus, the reference point for this goal is to have all assessed species in the region with an extinction risk status of Least Concern, scaled so that a score of zero is reached when 75% of species are extinct (following Halpern et al. 2012). Species distributions were determined using IUCN (IUCN 2017) and Aquamaps (Kaschner et al., 2015) species range maps. Species scores (Table A2-13) were averaged for each 0.5-degree cell, and then cell scores were averaged for each region (adjusting for the area of the raster cell and number of species present within the cell).

3.2.3.8 Artisanal Needs

This goal was altered from the original 'Artisanal Fishing Opportunities' following the approach taken by Daigle et al. (2017), who changed the goal to a more Canadian-centric approach. As such, this goal assesses what is required from the ocean to allow people to hunt and fish artisanally. I recognised three broad themes to assess this:

1) Shoreline sea ice extent – fluctuating and/or diminishing shoreline sea ice can physically restrict access for artisanal hunters and fishers and can shift species distributions, making them harder to track (Laidler et al. 2009; Huntington et al. 2016).

2) Extinction risk of artisanally targeted marine mammals – marine mammals are widely hunted across the Arctic for their furs, ivory and as a food source (Hovelsrud et al. 2008).
3) Sustainability of artisanally targeted fish stocks – to ensure that artisanal fishers have healthy fish stocks to harvest into the future (Zeller et al. 2011).

Shoreline sea ice scores (h) were calculated in the same way as the coastal protection goal, comparing the current condition of the previous three years (C_c) with a reference point of average extent 1979-2000 (C_r) so that:

$$h = \frac{C_c}{C_r}$$

(Eq. 4)

The less shoreline sea ice, the lower the score.

Extinction risk of targeted marine mammals (x_{mm}) was calculated in the same way as the Iconic Species sub-goal, but only included marine mammals that are artisanally targeted in each region (Table A2-12). The reference point was to have all targeted marine mammals at 'Least Concern' status:

$$x_{mm} = \frac{\sum_{i=EX}^{LC} S_i \times w_i}{\sum_{i=EX}^{LC} S_i}$$

(Eq. 5)

where for each IUCN threat category i, S_i is the number of assessed species and w_i is the status value (Table A2-13). Sustainability of artisanally targeted fish stocks (x_{art}) was calculated as per the Food Provision goal but included only those species listed as artisanally targeted in the Sea Around Us Project data, and with no under-harvesting penalty applied:

$$x_{\mathrm{art}} = \prod_{i=1}^{n} SS_i^{\left(\frac{C_i}{\sum C_i}\right)}$$

(Eq. 6)

where *i* is an individual taxon and *n* is the total number of taxa in the reported artisanal catch for each region throughout the time-series, and *C* is the average catch, since the first non null record, for each taxon within each region. Stock status scores (*SS*) are derived from B/B_{MSY} values – where for $B/B_{MSY} < 0.95$ (1.0 - 5% buffer), status declines with direct proportionality to the rate of decline of B with respect to B_{MSY}. No under-harvesting penalty was applied so any B/B_{MSY} score > 0.95, received a *SS* of 1.

The status for this goal is an average of scores for each of the sub-components:

$$AO = \frac{h + x_{mm} + x_{\mathrm{art}}}{3}$$

(Eq. 7)

Norway's score included only artisanal fish stocks, as artisanal marine mammal hunting is not practiced and sea ice is not plentiful or used for fishing.

3.2.3.9 Tourism and Recreation

A healthy ocean should provide tourism and recreation opportunities for people to enjoy. This goal uses employment in tourism as a proxy for the number of people engaged in tourism and recreation across the Arctic. As such it should respond dynamically to the number of people actively seeking tourist opportunities in each region, because if tourism increases or decreases then the number of jobs needed to service this sector should respond similarly. The number of tourism jobs are converted to percentage of employment in tourism to adjust for population size differences and multiplied by a sustainability coefficient drawn from the Travel and Tourism Competitiveness Report, which assesses countries for their overall quality, future potential and long term sustainability of tourism (Crotti & Mashri 2015). For a region-wide comparison, the reference point is set as the 90th percentile of the best performing region across all years, to account for outliers.

3.3 Results

Overall, the Arctic within EEZs scored 78 out of 100 (Figure 3-3). Scores varied substantially across assessed regions, from 65 in Jan Mayen to 87 in Svalbard (Figure 3; Figure S1). Averaged across the Arctic, three goals scored 90 or above: Biodiversity (95), Livelihoods and Economies (93), and Clean Waters (90). Marine Mammal Harvest scored 88 and Artisanal Needs 81. Coastal Protection scored 79 and Sense of Place scored 77, but there was disparity between the sub-goals, with Iconic Species scoring 85 and Protected Places scoring 68. Food Provision scored 67 overall, while its sub-goals of Mariculture and Fisheries scored 36 and 68 respectively. Tourism and Recreation was the lowest scoring goal (33).

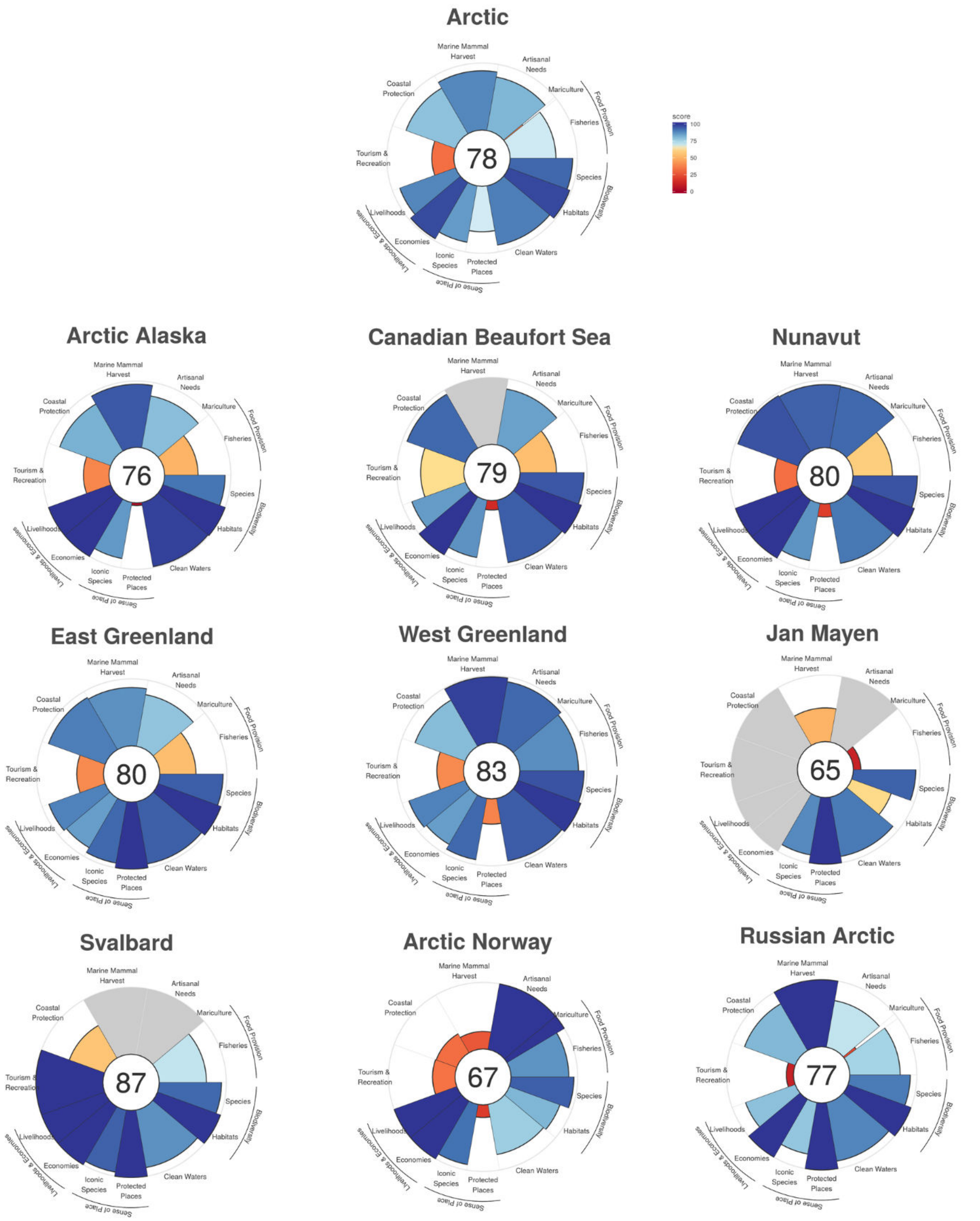

Figure 3-3: Arctic Ocean Health Index scores for each assessment region. Scores can range from 0 (bad) to 100 (excellent). Grey petals indicate that particular goal was not relevant to that region and thus not assessed.

Overall, species-related goals (Species sub-goal of Biodiversity and Iconic Species sub-goal of Sense of Place) scored highest across all regions but still show substantial room for improvement, with 82% of the 401 marine species assessed by the IUCN in the study region considered to be Least Concern, 5% Near Threatened, and 13% Threatened (Vulnerable, Endangered or Critically Endangered). Habitat-related goals (Coastal Protection goal and Habitat sub-goal of Biodiversity) presented mixed results. For soft-bottomed habitat, large areas remain free from disturbance of commercial fishing. For example, Arctic Alaska's northern coast is designated a Fishery Control Zone, with no commercial fishing activity allowed. As such, habitat scores were higher in regions which were more remote and under less fishing pressure (Russian Arctic, Greenland, Canada, Arctic Alaska). The Barents and Norwegian Seas are subject to trawling for key target species (Buhl-Mortensen et al. 2016), meaning Arctic Norway and Jan Mayen received lower scores for soft-bottom habitat (81 and 63 respectively). Sea ice related goals (sea ice edge extent within the Habitat sub-goal, shoreline sea ice extent for Coastal Protection) are also high with the exception of Arctic Norway and Svalbard, which are lower because the Barents Sea has experienced some of the most significant warming and variable sea ice conditions on the planet (Sato et al. 2014; Eriksen et al. 2017; Onarheim & Årthun 2017).

The Clean Waters goal scored well for all regions; this reflects the low population density of the Arctic in general and few sources of pollution. Norway scored the lowest (87), which was largely driven by a much higher chemical pollution score than other regions and reflects both the higher population density along the coastline and prominent shipping routes.

The Protected Places sub-goal of Sense of Place varied widely across regions, with both very high and very low scores. Although there has been an expansion of Arctic protected areas over the last 50 years, across the CAFF area, only 4.2% of marine areas are protected, with terrestrial areas garnering a much higher 20.2% protection, showing there is room for improvement in this area (CAFF & PAME 2017).

Marine Mammal Harvest scored high across all regions, except Arctic Norway and Jan Mayen, which scored lower primarily due to a significant under-harvest of seals compared to their quotas. Significantly, the Canadian commercial seal hunt occurs largely outside of the AOHI study area and so was not included. Marine mammals were also considered in the Artisanal Needs goal, which also included sustainability of artisanal fish stocks and extent of shoreline sea ice. Scores for this goal were generally high, in part because the fisheries component of the goal did not penalize underfishing and many artisanal stocks appear healthy.

Fisheries scores were between 50-75, with the exception of West Greenland (87), Arctic Norway (87) and Jan Mayen (11). High scores for Arctic Norway and West Greenland align with landings primarily being from Marine Stewardship Council (MSC) certified fisheries. The Jan Mayen score potentially reveals issues with the spatial distribution of the catch data from SAUP. For this region, 96% of landings are classified as “marine fish unidentified” – which is heavily penalised for poor taxonomic reporting in the OHI. Unidentified landings for Arctic Norway and Svalbard (part of the same management area) are much lower at 4% and 13% respectively, meaning the methods for distributing catch data for this region may be causing unfair penalty.

Running the Watson (2017) dataset through the AOHI reduced the overall Index Fisheries score by 13 (68 to 55), however overall AOHI scores declined by just one point from 78 to 77 when using this alternative fisheries dataset. From a regional perspective, Watson’s data resulted in a dramatic improvement of Jan Mayen’s Fisheries score, from 11 with the SAUP data to 59 with Watson’s (Figure 3-4). East Greenland also saw an improvement using Watson’s data (54 to 62). However, all other regions saw a decrease in scores. Arctic Alaska, Nunavut, Canadian Beaufort and Svalbard all had scores reduced by less than 10, yet Arctic Russia (20), Arctic Norway (28) and West Greenland (34) all had large decreases. . All regions except West Greenland (83 down to 79) and Jan Mayen (65 up to 75), showed a decrease in overall OHI scores of 2 or less.

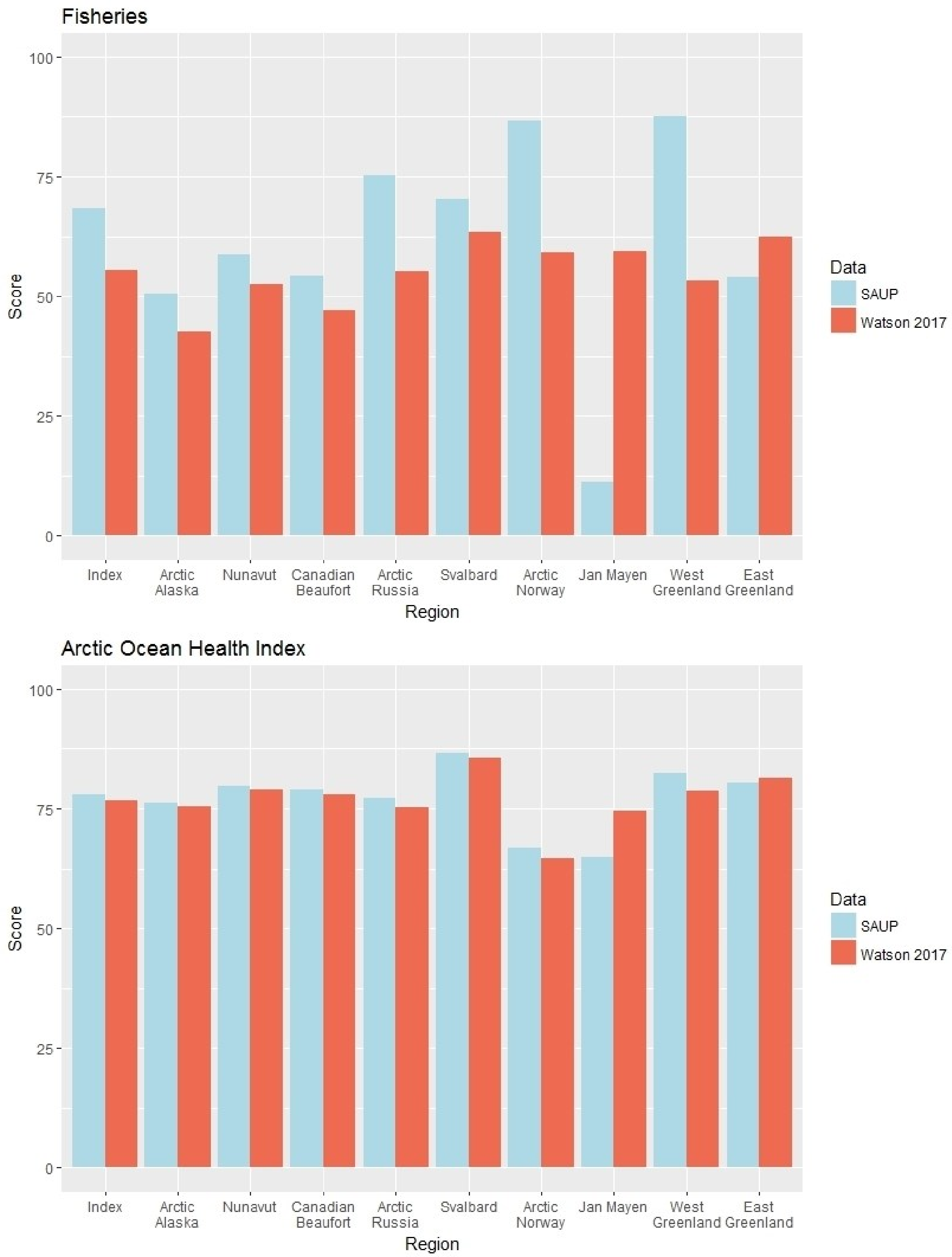

Figure 3-4: Comparing difference in scores for Fisheries (top) and overall AOHI (bottom) when using different fisheries catch data

3.4 Discussion

The Arctic is a globally unique and important geography for biophysical, cultural and economic reasons; yet its management is disjointed and resources for monitoring are limited. I have made the first attempt to piece together disparate data-sets across the pan-Arctic area to quantitatively assess ocean health from a human-centric perspective. Despite challenges, I have shown such studies are possible and provide an initial baseline of current pan-Arctic social-ecological conditions using freely available data, and sharing the AOHI framework and open code for future iteration and improvement (Lowndes et al. 2017). Given the Arctic is rapidly changing, baselines need to be established and systems continually evaluated in order to inform management (Hussey et al. 2016). I discuss the context of the results below and highlight spatial patterns of interest.

3.4.1 Spatial patterns and management considerations

Relatively high scores for species-related goals are likely due to rebuilding of once heavily exploited whale and pinniped populations, absence of large commercial fishing fleets in many parts of the Arctic, and increased productivity from climate change being beneficial to many fish species, at least in the short term (Mcrae et al. 2012). This assessment falls largely in line with the Arctic Species Trend Index (ASTI), which found an increase in Arctic marine vertebrates from 1970 to 2005 (Eamer et al. 2012). While the AOHI reflects the fact that in the short-term climate change may be having a potentially beneficial effect for many species, it does not account for long term risks to marine biodiversity. Arctic species often have particularly narrow temperature ranges and are highly susceptible to invasions. Projections indicate that the Arctic could be at high risk of invasive species and localised extinctions, highlighting the need for ongoing monitoring, and dynamic and predictive management (Cheung et al. 2009; Eamer et al. 2013; Garciá Molinos et al. 2016).

Arctic sea ice is critical for climate regulation, coastal protection and as habitat to a range of species that live in or around the ice or use it for life history events such as reproduction, moulting or resting (Laidre et al., 2015). It is therefore also important to people who depend on sea ice-associated species as resources and for cultural reasons. Sea ice dynamics in the Arctic are being altered by warming at twice the global mean rate, and some projections suggest an ice-free summer by 2040 (AMAP 2017). Therefore, measuring these changes is an important component of the AOHI (Overland & Wang 2013). High sea-ice related scores (Habitat and Coastal Protection) are in keeping with the global OHI assessment, which found reduced sea ice scores in sub-Arctic countries (Lithuania, Sweden, Finland, Norway [which includes Svalbard], Estonia, Latvia), but not yet at higher latitudes (Halpern et al. 2017). While the data I used considers temporal and spatial extent, it does not consider depth of sea ice, which may be important to sea ice-associated biodiversity (Kovacs et al. 2011). Further work could also include representing the shifting seasonality of sea ice (Haine & Martin 2017).

The Livelihoods and Economies goals do not currently consider informal economies or subsistence livelihoods, which are prevalent in many regions across the Arctic (Larsen et al. 2015). I was unable to find data to support the development of an indicator reflecting these less formal elements of the economy at the pan-Arctic scale, particularly as the informal economy varies widely between regions (Schmidt et al. 2015). Furthermore, many indigenous communities view economic development differently; while many wish to maintain traditional lifestyles, many communities are keen to mitigate high levels of poverty, ill health, and food security issues through full-time employment and the benefits that economic development can bring (Stewart et al. 2011; McCauley et al. 2016). These issues are inherently local, making it challenging to find a meaningful reference point at the pan-Arctic scale. Gaining a better contextualised understanding of how people conceptualise the elements of wellbeing within each region might yield information allowing relative change in wellbeing to be compared (Woodhouse et al. 2015).

Arctic tourism above the Arctic Circle is largely dominated by cruise ship tourism, which has grown markedly since 2008, particularly for Svalbard where tourism employment is high relative to the permanent population (Viken 2011). However, numbers remain far below more accessible sub-Arctic areas (Maher 2017), indicating that demand may be present and there is much room for growth, as indicated by the AOHI scores. The future of tourism in the Arctic is unclear but has the potential for significant social, economic, and ecological impacts, both positive and negative (Stewart et al. 2015). There are already concerns of exceeding carrying capacity in countries such as Iceland, which has seen a six-fold increase in tourism since 2008 (Maher 2017). Balancing the economic benefits of tourism while maintaining the environmental and cultural sense of place that makes tourism attractive is a difficult undertaking. Setting out a shared vision for Arctic tourism and developing infrastructure in areas that will have positive social impacts and minimise negative environmental impacts should be a priority for Arctic nations.

The Marine Mammal Harvest goal indicates that marine mammals with population data are being harvested sustainably across the Arctic, showing that sustainable management of these species is possible if supported by scientific research. For example, although controversial, the management of the Canadian harp seal hunt can be considered a conservation success, with the number of individuals rising from a low of 1.1 million in the early 1970s to over 7 million today. This supports the importance of robust monitoring and evidence-based quotas (Hammill et al. 2015). However, the analysis of Marine Mammal Harvest was restricted by the need for information relating to both landings of marine mammals and viable population estimates leading to quotas or PBR estimates, which unfortunately excluded most marine mammals in each region. Abundance and trend data for Arctic marine mammals is poor or largely absent, which makes quantifying the sustainability of harvests difficult. Obtaining population estimates for new metrics from sources such as indigenous knowledge or the Circumpolar Biodiversity Monitoring Plan (Gill et al. 2011) should be a priority in aiding management of marine mammals, particularly in the face of climate change (Laidre et al., 2015; Gill et al. 2011). Ideally, establishing metrics and quotas would be a systematic process that would enhance co-management of marine mammal species locally and at a pan-Arctic scale.

The Food Provision scores reward areas with the highest levels of sustainable catch, showing the benefits of well managed and productive fisheries. Arctic fisheries are an increasingly controversial topic; seemingly offering large potential for food provision and economic benefits, but with high ecological risk (Lam et al. 2016). The USA has recognised this risk by closing a large proportion of its Arctic EEZ to industrial fishing and all the Arctic countries have signed an agreement to prohibit fishing in the Central Arctic Ocean. Fisheries in the Barents and Norwegian Seas have been recovering from previous exploitation, improved management and beneficial effects of climate change, which have led to a threefold increase in spawning stock biomass in the last 15 years (Dalpadado et al. 2014; Grønnevet 2016). Using a different dataset revealed changes in scores, showing that data selection can be a critical component driving scores. While this study often did not have the luxury of multiple datasets, exploring two Fisheries goal datasets revealed potential issues in each, showing that this type of sensitivity analysis would be useful in similar studies. The increased score for Jan Mayen with Watson's data shows that key differences exist in the spatial disaggregation of the catch data in this area; the higher SAUP scores for Norway and West Greenland, where commercial fisheries operate mainly under MSC certification, align with what I would expect to see.

3.4.2 Implications for future pan-Arctic management

The AOHI assessment provides a starting point for consideration of pan-Arctic social-ecological dynamics, which like other composite indicators can be iteratively improved over time as more and better data become available and dynamics are better understood (Burgass et al. 2017). Understanding current limitations and how social-ecological systems are changing is necessary for effective management (Harris et al. 2017). Many of the goals within the AOHI are transboundary in nature and require co-management (Biodiversity, Marine

Mammal Harvest, Coastal Protection, Tourism and Recreation, Clean Waters, and Fisheries). Consideration of these interlinkages is critical for management; the heterogeneity of the Arctic means that system dynamics are important not only between different goals but also across regions and localities. For example, co-management of many marine mammal species such as walrus and polar bears across national borders and between indigenous groups has been successful in ensuring sustainability and maintaining human wellbeing (Laidre et al., 2015).

The AOHI provides a snapshot in time of the current status of the Arctic, as well as an indication of the near-term future state. The biggest driver of change in the Arctic is climate change. Although it is included in the assessment and its effects are already being felt, these will be most noticeable over the medium and long terms (Bennett et al. 2015). Subsequent AOHI assessments will be required to track the impacts of climate change across the region. As the Arctic ‘opens up’, the opportunities for economic development will become even more numerous for oil and gas extraction, tourism, shipping and infrastructure. Understanding the risks that climate change poses to the health of the ocean, and the wellbeing of the people who depend on ocean resources, will require multi-faceted modelling, with a strong emphasis on social science (Ford et al. 2015). Given the sensitivities of people and environments in the Arctic, pan-Arctic assessments such as ours can help inform decision making on strategies for investment to minimise social-ecological risk and maximise benefits across the region. Pan-Arctic plans for environmental protection and sustainable development would limit ad-hoc developments, which could otherwise pose severe risks to unique ecological communities or areas of biodiversity. Similarly, given the heterogeneity of the region and in order to protect and restore the full range of biodiversity across the Arctic, coordinating efforts across large scales is required to ensure an ecologically coherent network of protected areas (Harris et al. 2017).

A key area of uncertainty for marine management in the Arctic is the potential for tipping points or thresholds, which can be classified as periods of rapid, non-linear change (Serrao-

Neumann et al. 2016). Given the range and scale of pressures on the Arctic, particularly climate change, these tipping points may well be crossed even in the short-term, which would compromise the predictions of the AOHI. Late action to halt or reverse a tipping point is highly ineffective compared with early identification and preservation of system resilience (Selkoe et al. 2015). Through this assessment I provide data for a range of different physical, biological and social data layers that can be accessed from raw data through to aggregated final scores of the AOHI to inform system-wide management. Monitoring of these layers should be focused towards identifying potential tipping points in order that pre-emptive action can be taken.

Ultimately, threats such as climate change go beyond pan-Arctic governance and will require global mitigation in reducing greenhouse gas emissions. Despite climate change being a huge risk to the Arctic, the disconnection of communities, authorities and governments at the pan-Arctic scale prohibits a clear and united message. This work provides foundational datasets which can be of use for both pan-Arctic assessment and local-decision making for Arctic futures. Ensuring a participatory process and inclusion of the full range of stakeholders is often vital in ensuring evaluation of management strategies or alternative futures is appropriate and useful (Dichmont & Fulton 2017). Promoting pan-Arctic monitoring, management and decision-making, joined with a bottom-up approach of case studies and storytelling, could help position the Arctic as a bellwether for climate change and help create increasingly ambitious, robust and equitable climate policy at the global scale. A data-driven approach such as the OHI allows the quantification and clear communication of broad results to a range of stakeholders both inside and outside the Arctic. It can therefore help communicate complex issues and include more stakeholders through transparency and open web-based tools. However, when being used for management purposes, the data and models must be interpreted carefully as with any scientific output. As such, I have been transparent with the data that has been used and their limitations such that these can be factored in to any future use of the AOHI.

The OHI goals are intended to provide a broad comparative framework, and to encourage thinking about what we consider to be a "healthy" system, and how far we are from that state in relation to a range of human-defined functions and goals. It can be updated and improved iteratively over time should new data become available, or should new social considerations need to be factored in. This will be particularly important for risks such as ocean acidification and climate change, which has the potential to alter the structure and biodiversity of ecosystems (and subsequently impact people), but the long-term effects of which are not well understood (Lam et al. 2016). Likewise, while oil and gas extraction is not included here (as with other OHI assessments), due to it being intrinsically unsustainable, it is undoubtedly of huge economic and social importance to some Arctic areas and therefore stakeholders may wish to include it in future assessments.

3.5 Conclusions

The AOHI is a first step towards measuring the status of the ocean across the high pan-Arctic area. In general, I found the Arctic to be sustainably delivering a range of benefits to people, with room for improvement in all goals, but particularly in sustainable tourism, mariculture, fisheries, and protected places. Biodiversity-focussed goals presented encouraging scores, showing how improved ecosystem management through recovering fisheries and sustainable marine mammal exploitation were having a positive effect. However, the assessment was constrained by limitations in pan-Arctic data, in particular the disjointed and non-comparable nature of data from different Arctic regions. While validating a composite index is a difficult undertaking, the process of its formulation and understanding where conceptual and data uncertainties are located is inevitably crucial for informing management. Obtaining comparable data from across the Arctic to minimise these uncertainties is a priority for informing robust pan-Arctic stewardship; such efforts should be targeted towards the most pressing and urgent transboundary management challenges, such as fisheries, biodiversity and economic development (shipping, tourism, extractive activities; Tesar et al., 2016).

4 VALIDATION AND USE OF LARGE-SCALE BIODIVERSITY INDICATORS AT THE NATIONAL SCALE

4.1 Introduction

Tracking the status and fate of biodiversity remains a fundamental conservation and management need, and challenge. Many international conventions and treaties have set forth goals to halt (and ideally reverse) biodiversity declines and have established targets to help managers make progress towards these goals. For example, the Convention on Biological Diversity's (CBD) Aichi Targets provide global aspirations for biodiversity. Yet the degree to which global biodiversity indicators can adequately track progress towards these targets remains unclear (Tittensor et al. 2014). In part these shortcomings are due to the targets being developed largely without consideration of whether and how they can be quantified by specific indicators (Butchart et al. 2016), but also because many indicators suffer from a variety of uncertainties related to their mathematical construction or the data that underpins them (Burgass et al. 2017). There are several desirable characteristics of the global biodiversity indicator suite: they should be cost-effective, taxonomically diverse, frequently reported, meaningful to the public, informative across scales, should reliably inform status and trends of biodiversity, and respond predictably to policy changes (Jones et al. 2011). Indicators based on species metrics, such as abundance, often struggle to meet these characteristics (Stephenson et al. 2017) and therefore require additional validation that many are not currently subject to (Moriarty et al. 2018). Global species-based biodiversity indicators, such as the Red List Index or the Living Planet Index, utilise species data from across the world and synthesise information at the planetary scale. These large-scale indicators tend to be cost-effective by utilising open source data, are reported upon frequently and resonate well with the public. However, they have been criticised for having limited taxonomic diversity (with bias towards well studied species and certain geographies) and little or no in situ validation; thus it is often unknown how well they reliably inform users about status and trends of biodiversity (Collen & Nicholson 2014).

An important consideration when selecting global biodiversity indicators was their applicability across scales, particularly for national-level decisions (Biodiversity Indicators Partnership 2018). The Ocean Health Index has made this central to its design, providing data and scores for each country in the world (Halpern et al. 2012). However, the limitations around taxonomic bias and lack of validation has meant global indicators are rarely disaggregated and used for decision-making at the national scale. Nations have therefore tended to use their own existing biodiversity indicators or created new ones (i.e. Government of Pakistan (2017) set their own indicator approach for monitoring biodiversity goals in their National Biodiversity Strategy Action Plan), rather than improving the science and uptake of the existing indicator suite, which is already linked to high-level political goals. While this approach is more straightforward for countries, often easily fitting in with ongoing monitoring programmes, the national indicators that are used are still likely to have the same uncertainties around reliability and the predictability of their response to change, and are often not subject to validation. There is therefore an urgent requirement for further validation of species-based biodiversity indicators across all scales in order to refine indicator suites (Moriarty et al. 2018). Here I highlight why a lack of consensus on biodiversity indicators is problematic across scales and demonstrate how indicator validation using a modelling framework at the national scale can be beneficial to understanding indicator performance and achieving agreement on species-based biodiversity indicators.

4.1.1 Connecting biodiversity indicators across scales

Current global biodiversity indicators draw on synthesised global datasets to report on progress towards international biodiversity goals. This remains largely separate from the CBD process, where the achievement of biodiversity goals is predicated on the commitments and actions of member states who pledge and report on their progress through National

Biodiversity Strategies and Action Plans (NBSAPs) and National Reports (NRs) respectively. Synthesis of these NRs is carried out by the CBD, who assess how well aligned national commitments are to the Aichi Targets and how much progress is being made towards them by each country (and in turn globally; Convention on Biological Diversity 2018). Data that are used to generate local or national indicators and reported through the NRs is therefore often not feeding into the global indicator suite, which would help to improve the overall global taxonomic and geographic diversity of global indicators. There is therefore a disconnection between monitoring, reporting and decision-making at the national and international scales; this means that global indicators may not be representative of global biodiversity and local or national indicators are poorly understood and have difficulties in scaling up to the global level. These issues hinder the ability of indicators to be used for policy and scenario testing. If indicator validation was improved, a consensus around the appropriate indicators to use to monitor biodiversity change could be achieved, thereby focussing monitoring efforts and enabling nations to use data in a standardised way to global effect (Figure 4-1). While the period of the Aichi Targets is coming to an end, the post-2020 biodiversity agenda will probably see a new set of targets formulated. Effective science-based targets will require clear leadership on how they are achieved through an analysis of policy pathways for biodiversity; in order to achieve this there will need to be further development and testing of biodiversity indicators across a range of scales.

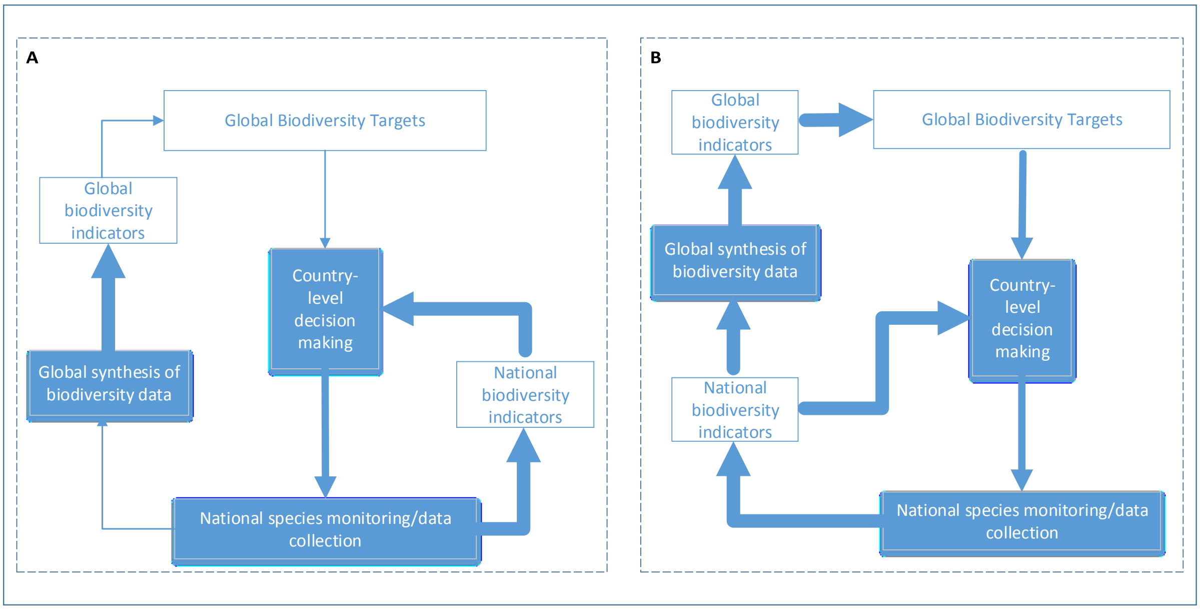

Figure 4-1: Disconnection and reconnection of global and national biodiversity indicators. Width of blue arrows show strength of connection. A) Current disconnection of national biodiversity monitoring and indicators, which are poorly linked back in to global biodiversity indicators and targets. Subsequently this means that global targets are more poorly formulated and have a weaker connection to national level decision making. In B) there is consensus between national and global biodiversity indicators, which increases the overall quality of global biodiversity indicators being fed from a variety of national sources. This allows for improved global target formulation and thus a stronger connection to country-level decision making.

4.1.2 Models and indicator testing for consensus

Achieving consensus on a set of biodiversity indicators which clearly meet the criteria set out by Jones et al. (2011) would allow for biodiversity to be better linked to policy making, by enabling users to project scenarios and understand the potential effects of policy interventions on biodiversity across a range of scales (Pereira et al. 2013). However, consensus can only be achieved if there is confidence in the ability of indicators to be useful proxies for biodiversity change. Fisheries science has focussed much attention on the development and testing of indicators which report on the effects of fishing (Fulton et al. 2005), given the strong and direct link between fishing mortality and fish species status. This has enabled fisheries managers to gain confidence in a suite of indicators that point towards underlying changes in the ecosystem; this allows for more robust target setting, quantitative measurement of target achievement and analysis of policy scenarios (Fay et al. 2013). Global biodiversity indicators have received limited testing, which means that they have mostly been used to report on trends and remain poorly linked to overall target setting and policy evaluation (Collen & Nicholson 2014).

Understanding how indicators respond based on an analysis of real-world data is challenging as biodiversity is complex and can be prone to non-linearities and driven by numerous factors. Indicator testing is therefore best done using models, as they allow for explicit consideration of the drivers underlying change in the indicator, look at the effects of different policies and thus understand the link between intended actions and outcomes. Using quantitative models to generate indicators through various scenarios can thus reveal system dynamics that might cause unintended responses in indicators from unanticipated responses to policy (Nicholson et al. 2012) or reveal issues with indicator construction (Costelloe et al. 2015). As such, models can be used to mimic the real world and gain a virtual 'truth' not possible in real world datasets (Branch et al. 2010). This helps users to understand whether or not indicators respond predictably to interventions and tease apart different drivers of change to inform decisions (Link et al. 2010).

Many new indicator frameworks have been proposed since 2010 that have found use at the national scale, with the potential to scale up or feed into a global system (Certain et al. 2011; Pereira et al. 2013; Coll et al. 2016; Miloslavich et al. 2018). Even with perfect representation of species, without understanding the drivers of change that underlie the data, how they relate to each other and how change in the world manifests as change in the indicator, then indicators will remain of limited use to decision makers (Collen & Nicholson 2014). In order to obtain the most useful indicators for reporting against science-based targets, it is necessary to test a wide variety of indicators across different models to understand how they respond to changes within the system. This would help to transition from indicators simply reporting on trends to being useful tools to assist with decision making. Here, I demonstrate how a modelling framework can be used to support biodiversity indicator validation through exploring indicators' responses to system changes.

I used the Norway and Barents Sea Atlantis model (NoBa) to generate two different types of species-based biodiversity indicator; the Living Planet Index (LPI), a global indicator, and the Norway Nature Index (NNI), a national indicator, representing two different types of indicator construction method, based on abundance data, that have received little testing. I parameterised the model to an unfished system to ask the question does sustainable fishing or overfishing lead to damaging trends in the wider ecosystem? I compared these two scenarios to a no fishing scenario over a period of 35 years to see how each indicator responded to each management scenario. Fisheries have some of the largest impacts on marine systems (Halpern et al. 2015a) and by simulating varying fishing pressures on an unfished system, I aimed to perturb the virtual ecosystem and trigger indicator responses. I analysed the extent to which the indicators can distinguish between the scenarios and if they could report conclusively whether fishing was damaging. I then evaluate the extent to which model-based testing gave insights into indicator behaviour.

4.2 Methods

4.2.1 Model Overview

The Norway and Barents Sea (NoBa) Atlantis model was developed by the Institute of Marine Research in Norway to represent the key species and processes in the Nordic (Norwegian, Greenland and Iceland seas) and Barents Seas, with the intent of exploring combined climatic and fisheries scenarios (Hansen et al. 2016).

Atlantis is an end-to-end marine ecosystem model that provides a repeatable and transparent basis for modelling ecosystem dynamics, considering different parts of marine ecosystems - oceanographical, biological, economic and social (Fulton et al. 2011a). The need for such models has been amplified by the desire for ecosystem-based management. Originally focused on ecology and then fisheries dynamics, it has begun to be used to assess interactions among and impacts of multiple uses of marine systems beyond fisheries, as well as climate impact questions. Full details of the Atlantis model, including the user guide, can be found in Audzijonyte et al. (2019); here I summarise its key components and setup. At its core it consists of deterministic physics and ecology submodels, which are spatially-resolved in three dimensions using a map made up of user-defined boxes and depth layers, as well as three types of habitat; water column, epibenthic habitat and sediment. The physics submodel includes oceanographic processes such as water fluxes, salinity and temperature which are often forced from relevant oceanographic models. Together with forcing of nutrient inputs, oceanographic processes drive deterministic primary productivity and influence movement of organisms between polygons. The ecology submodel explicitly tracks the flow of nutrients through trophic levels, with the main ecological processes being production, consumption and predation, waste production and cycling, migration, reproduction and recruitment, habitat dependency and mortality. Species are represented either by biomass pools or age-structured groups. Biomass pools typically are used for invertebrates, while age-structured groups used for vertebrates. Age-structured groups are modelled using the principles of physiologically structured models, where growth and reproduction is dependent on food availability and feeding interactions and reproductive output depend on the realised size and condition. For age-structured groups, all individuals within an age group are identical within one cell of the model domain, and their numbers and body condition are tracked as energy allocation to structural and reserve nitrogen pools.

Fishing fleet dynamics are included through the harvest submodel and can be standalone or used with an economics submodel but can be customised to simulate management or exploitation regimes. At its simplest this can include constant fishing on a single functional group or at its most complex dynamic fisheries comprised of interacting fleets with different gear, bycatch and management characteristics that are affected by compliance decisions and economic incentives such as quotas or fish price.

Atlantis has been applied in assessing alternative fishery management strategies, historical impacts of harvesting, compliance with fishery regulations, robustness of ecological indicators, impacts of global change (including coral), effects of changes in fish body size on ecosystem dynamics, and the implications of model complexity. Given the overall Atlantis model is a framework, validation is required for individual model applications in different parts of the world. The extent to which this has been done differs due to the complexity of the model and complexity of model validation in general, but examples exist and the field of end-to-end model validation is growing (Olsen et al. 2016; Ortega-Cisneros et al. 2017; Hansen et al. 2019; McGregor et al. 2019).

As with all Atlantis models, NoBa is a spatial box model, covering the area shown in Figure 4-2. The total area is 4 million km^2 divided into 60 polygons, which were decided upon by a group of experts covering fields such as oceanography, demersal fish, pelagic fish, benthos and marine mammals. Polygons were created to be relatively homogenous with respect to hydrography and bathymetry, as these are important features determining the distribution of biota in the Barents Sea. The boundaries of the model are in large part defined by "natural" boundaries, such as land and topography. Each polygon has up to seven depth levels, depending on their total depth. The depths of the vertical layers are 0-50m, 50-150m, 150-250m, 250-375m, 375-500m, 500-1000m and 1000-1200m. If the mean depth of the polygon is more than 1200 m, the lowest depth level will stretch to the bottom.

The NoBa model was initiated from 1981. Atlantis requires time series of temperature, salinity and volume fluxes across the polygon borders, and NoBa gets these from a Regional Ocean Modelling System (ROMS; Shchepetkin & McWilliams 2005) covering the Northeast Atlantic (Skogen et al, 2007). Within the model, 53 different functional groups represent the ecosystem, including detritus, carrion, bacteria, zooplankton through to marine mammals such as polar bears and killer whales (see Appendix 3 for full list). Many functional groups are individual species, but some groups are aggregations of several species.

In order to perturb the system in a way that we would expect the biodiversity indicators to pick up, I ran the model under three different fishing scenarios for seven different commercial fisheries consecutively; $F_{MSY}0$ (no fishing), $F_{MSY}1$ (considered a desirable rate of fishing in the Barents Sea) and $F_{MSY}2$ (overfishing). $F_{MSY}1$ is the maximum rate of fishing mortality that will eventually result in a population size of B_{MSY}. B_{MSY} is the biomass that enables a fish stock to deliver the maximum sustainable yield and is often used as a reference point for managing fisheries. The baseline rate of fishing for these runs was no fishing. This meant that for the $F_{MSY}1$ and $F_{MSY}2$ scenarios they were initially fishing an unimpacted system.

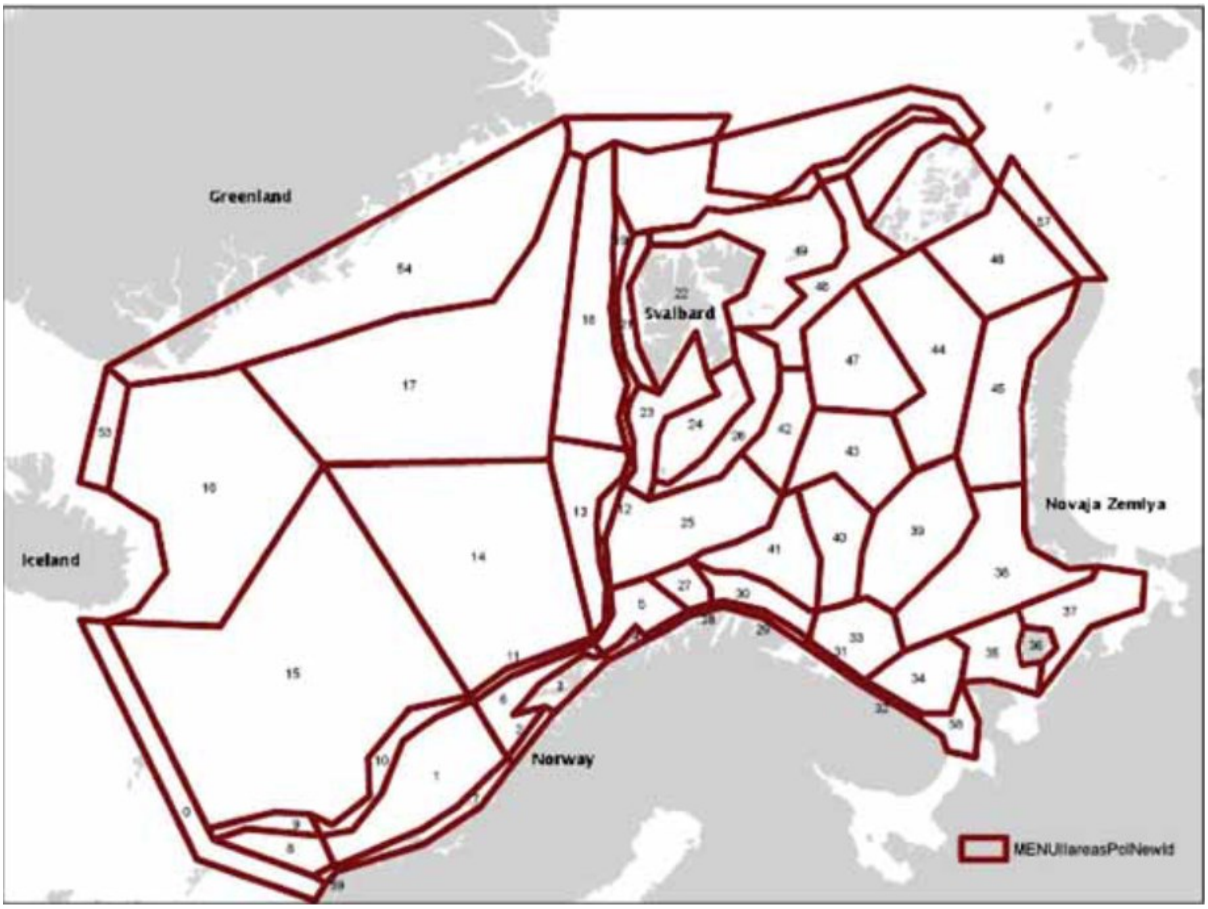

Figure 4-2: Model domain for the Norway and Barents Sea Atlantis Model

4.2.2 Indicator overview

The Living Planet Index (LPI) is a composite of time series of vertebrate abundance and biomass (Collen et al. 2009). It was adopted in 2006 as a headline indicator to the Convention on Biological Diversity (CBD) and is now used to track progress towards Aichi Target 12 (UNEP 2006). It contains trends for 14,152 populations of 3,706 species from across marine, terrestrial and freshwater realms and has had its methods and data continuously updated since its inception in 1998 (McRae et al. 2017). While the LPI has been disaggregated to track trends in regions such as the Arctic (Mcrae et al. 2012) and for species classes such as reptiles (Saha et al. 2018), it is rarely disaggregated for use at the country scale, a key requirement of CBD indicators. A disaggregation was performed for the Netherlands, where the globally available data were heavily supplemented with local data not included within the global LPI database, including from non-vertebrate groups (van Strien et al. 2016). As a response to the taxonomic biases that are inherently introduced from utilising available data, the LPI has altered its approach to weighting, to make the index more representative of vertebrate biodiversity (McRae et al. 2017), although invertebrates are still not included due to the lack of availability of consistent time-series.

The Norway Nature Index (NNI) is built from the Nature Index (NI) framework, which was proposed to facilitate the transfer of information from science toward other areas of society (Certain et al. 2011). The NNI was designed to show trends in biodiversity in major ecosystems across both marine and terrestrial realms. It is based on a large number of indicators representing different aspects of biodiversity. The overall objective is to measure whether Norway is succeeding in halting the loss of biodiversity, as has been pledged under several international agreements, but mostly notably under the CBD. The Norwegian Government made the Norwegian Environment Agency responsible for developing a biodiversity index to document overall trends for major ecosystems and the species they support and has committed to updating it every five years (Norway Ministry of Climate and Environment 2016). The NNI is included in the Norwegian official set of indicators for sustainable development, presented annually in the reporting on sustainable development indicators by Statistics Norway and by the Ministry of Finance in the National Budget.

The NNI aims to give an overview of the state of Norway's environment by focussing on trends in major ecosystems including forest, mires and wetlands, open lowland, fresh water, coastal waters, and the open sea. For each of these ecosystems, indicators were chosen for a variety of species groups, so that they aimed to be representative of overall species diversity. Indicators were selected from the main species groups – algae, lichens, fungi, plants, invertebrates, fish, amphibians, birds and mammals. In addition, indirect indicators that give information on the biodiversity potential of an area were included, for example the presence of dead wood and the degree to which open lowland landscapes (semi-natural habitats) are becoming overgrown. In all, the NNI uses more than 300 indicators in totality.

The NNI is a composite index, which is an amalgamation of indicators that are originally measured on different scales, but normalised to allow for comparison and aggregation (Becker et al. 2017). It uses expert judgement to selects indicators which are taxonomically representative, represent different trophic levels and functions of species, include common, rare and key species, are collectively sensitive to different types of pressures and represent different habitats within major ecosystems (Pedersen et al. 2016).

As such, the data the NI seeks to encompass and the structure of the index itself are selected in order to obtain meaningful ecological representation (e.g. covering a range of species that have different functional roles) and present the information in a way that is transparent and accessible to policy makers and environmental managers (Aslaksen et al. 2015). The NI was designed for use at the national scale and therefore does not aggregate up to the global biodiversity targets such as the CBD, although this could be possible with widespread uptake (Certain et al. 2011). The NI framework utilises a reference point approach to normalising indicators within the index, which is common for this kind of indicator (Burgass et al. 2017).

This means a reference point for a desirable state is established for each indicator and any value over this reference point is regarded as fully desirable and scored as ‘1’. Indicator values below the reference point are scaled to a score of ‘0’. This does, however, present problems for global aggregation if desirability is differently defined between countries.

The NI and the LPI are designed based on different approaches to biodiversity monitoring and reporting (see Figure 4-3). The LPI synthesises disparate abundance data of vertebrates from around the world in a consistent way for use at the global scale, but can subsequently be disaggregated by taxonomy, ecosystem type or region. The NI uses species data across several different ecosystems in an attempt to factor in ecological representation and selects sub-indicators to be representative across taxonomic groups including fish, mammals, birds, invertebrates and plants as well as applying weightings for ecological function. Both indicators in essence utilise the same types of abundance monitoring data, but for different species and in different ways.

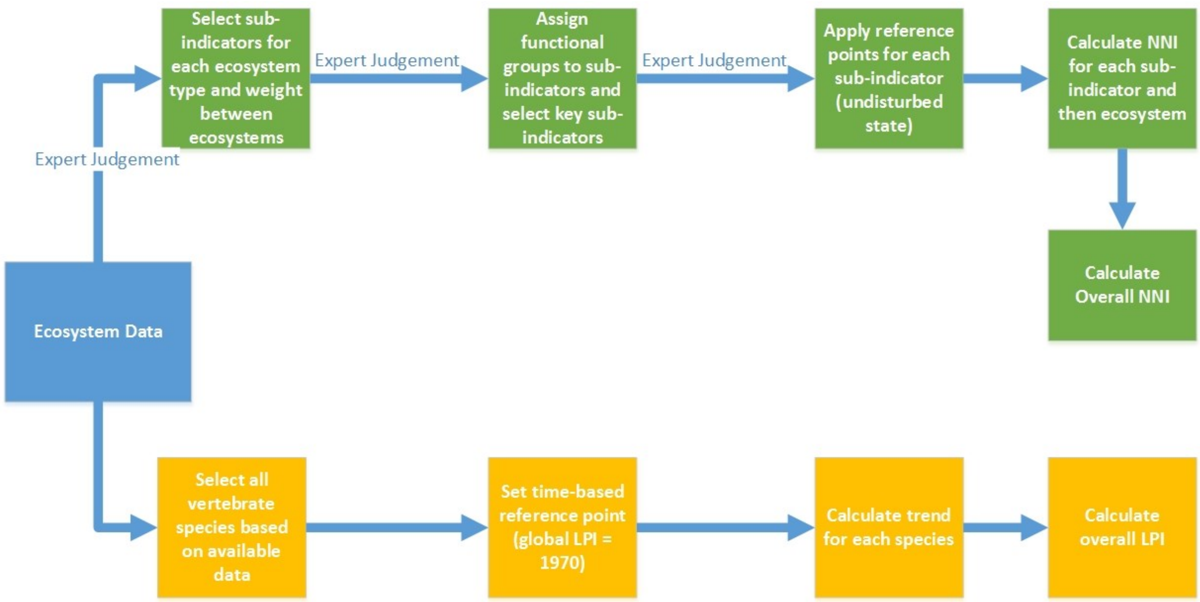

Figure 4-3: Steps taken to produce Norway Nature Index (green boxes) and Living Planet Index (yellow boxes)

I used the NoBa model to generate the LPI and NNI for the Norway and Barents Seas across the whole model domain, with the intention to see how changes in the system are reflected in these indicators. The LPI was generated using the 47 vertebrate species contained within the NoBa model (regardless of whether or not they were included in the LPI database), representing 29 functional groups. The LPI uses relative measures of change in abundance of species as input data; I utilised modelled biomass and assumed changes in species biomass as a result of the management scenarios were proportional to changes in abundance, as per Nicholson et al. (2012). I calculated the LPI using the 'R' package, *rlpi* (Zoological Society of London 2018)*,* calculating a geometric mean of trends for each species using Generalised Additive Modelling (GAM; (Collen et al. 2009). If the GAM fit for a particular species was poor, modelling was conducted using the chain method (Loh et al. 2005). Each species within the model was considered part of a single population, trends in which were calculated using the logarithm of the ratio for successive years (*d*)

$$d_t = \log_{10}\left(\frac{N_t}{N_{t-1}}\right)$$

(Eq. 1)

Where *N* is the species population and *t* is the year. The index value (*I*) is subsequently calculated in year *t* as

$$I_t = I_{t-1}10^{dt}$$

(Eq.2)

with the index value set to 1 in year 1. This treats all species as equally weighted; I did not pursue taxonomic weighting in this study as the species included were deemed to be representative of the ecosystem and such weighting can be considered subjective. A bootstrap resampling technique was then applied for 10,000 iterations in order to generate 95% confidence limits. Such an approach demonstrates the uncertainty in the index inherited from the baseline in year one and propagated through the time series, not that there is no uncertainty associated with year one (Collen et al. 2009).

I calculated the NNI for the demersal ocean and pelagic ocean ecosystems. Coastal ecosystems could not be included separately as the NoBa model was not designed specifically for coastal ecosystems, although coastal areas fall within the model domain. Sub-indicators (which are the selected species) from the NNI which were represented by functional groups in the NoBa model were included in the analysis. I generated 55 species sub-indicators using biomass trends from 24 functional groups within the model. Indicators were scaled between 0 and 1 using a reference value and a non-linear scaling function. The NNI reference point should be equivalent to an intact natural environment, with little human activity and as such I used the modelled virgin biomass for each sub-indicator. All sub-indicators were normalised using the LOW scaling model, meaning there is a positive relationship between these indicators and biodiversity (Pedersen et al. 2016):

$$S_t = \begin{cases} \frac{U_t}{U^{ref}}, & 0 \leq U_t \leq U^{ref} \\ 1, & U_t > U^{ref} \end{cases}$$

(Eq. 3)

where S is the normalised indicator, U_t is the indicator before scaling, t is the year and U^{ref} is the indicator's reference value. Weights for each sub-indicator take into account the sub-indicator's specificity to a given major ecosystem (specificity weight) as well as the sub-indicator's ecological function (trophic weight). Sub-indicators were assigned to one of six functional groups (top predator specialist, top predator generalist, intermediate predator specialist, intermediate predator generalist, herbivore/filter feeder, primary producer), which were given equal weighting within the index. Some species are considered key sub-indicators if they are representative of populations of a hundred species or more, occur in a large area and are well documented with good data; key sub-indicators were given additional weighting, accounting for 50% of the index (Certain et al. 2011). Specificity is the extent to which the indicator reflects the major ecosystems. As this study included only two major ecosystems, species could belong to either the pelagic ocean and/or demersal ocean ecosystems, with each indicator's total specificity being 100% (i.e. Greenland halibut was assigned 60% benthic and 40% pelagic). Weighting and assignment of functional groups matched that of the NNI wherever possible. Overall weights were determined as a product of both the trophic weight and specificity weight (Pedersen et al. 2016). The NNI (*NI*) was thus calculated for each major ecosystem (*j*), for each year (*t*):

$$NI_{jt} = \sum\nolimits_{i=1}^{n_j} S_{ijt}\, w_{ij}$$

(Eq. 4)

Where w is the weighting term and summation is over indicators documented from the major ecosystem.

The LPI and NNI were generated for each year of the model run for each of the three scenarios ($F_{MSY}0$, $F_{MSY}1$, $F_{MSY}2$).

Regression analysis

To analyse whether the indicators distinguished between management scenarios I used linear regression to fit a slope coefficient. It was hypothesised that a useful indicator should be able to differentiate between scenarios such that: lm(indicator ~ year) would be statistically significant versus lm(indicator ~ year + scenario) when using an analysis of variance test.

Regressions were subsequently visualised by using the regression equations to plot trends for each indicator and scenario with 95% confidence intervals. Secondary exploration was conducted by eye to see if the trend of the indicator reflected the expectation for a given scenario; it would be expected that biodiversity would decline more strongly under $F_{MSY}2$ than $F_{MSY}1$ than $F_{MSY}0$.

4.3 Results

4.3.1 Ecosystem dynamics

As was expected, demersal fish declined most severely under fishing pressure, as their biomass was directly removed from the system (Figure 4-4). With the exception of capelin (*Mallotus villosus*), all commercial fishing species' biomass declined most sharply and consistently under $F_{MSY}2$, while remaining largely stable under the no fishing scenario ($F_{MSY}0$). $F_{MSY}1$ saw a similar, but lesser, decline to $F_{MSY}2$ before both fishing scenarios stabilised. Capelin exhibited unusual patterns of a large spike in its biomass under the heaviest fishing pressure near the start of the scenario before a steep decline and similar, but lesser spikes under $F_{MSY}1$.

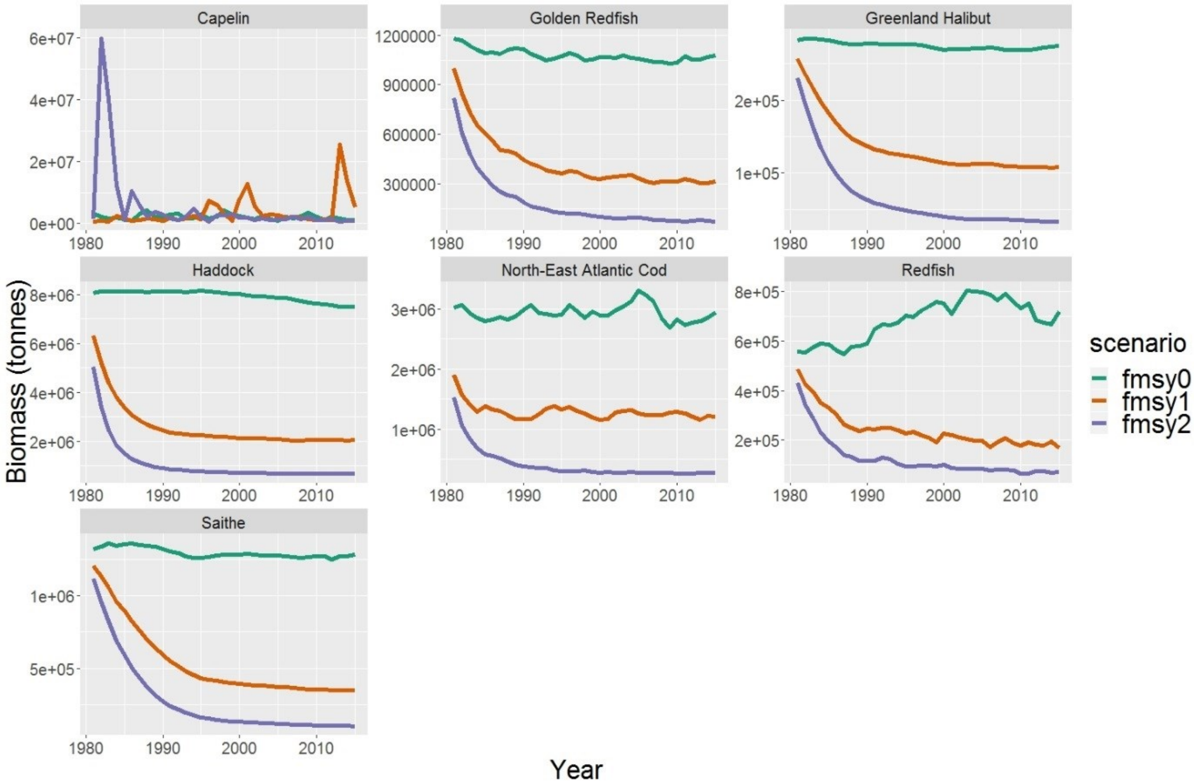

Figure 4-4: Changes in biomass of commercial fish species exposed to three different fishing pressures in the NoBa Atlantis model.

Diverse changes were seen in the wider ecosystem when comparing each of the three management scenarios (Figure 4-5, A). Changes in functional groups were most prevalent for those groups that were highly reliant on or predated by these demersal fish species; assemblages such as sea birds and marine mammals which have large ranges and varied diets showed little difference between scenarios (biomass changes for each functional group can be found in Appendix 3). Assemblages overall showed both gains and losses, showing the importance of understanding system dynamics. There were differing responses of functional groups within assemblages (Figure 4-5, B); notably pelagic fish and invertebrates showed large variation. One species of pelagic fish, Capelin, had a particularly severe increase under the $F_{MSY}1$ scenario, increasing by over 600%. Capelin have an extremely high reproductive potential and their stock size can fluctuate enormously over time as a result of both ecological system feedbacks and exploitation (Hjermann et al. 2004). Its population increase in this case is likely to be the result of the species itself coming under heavy fishing pressure, while many of its predators were also being removed from the system, allowing for great reproductive expansion. Invertebrates (such as prawns and cephalopods) experienced high inter-annual variation but overall increased their biomass under increasing fishing pressure, due to being released from predation pressure by reductions in demersal fish populations.

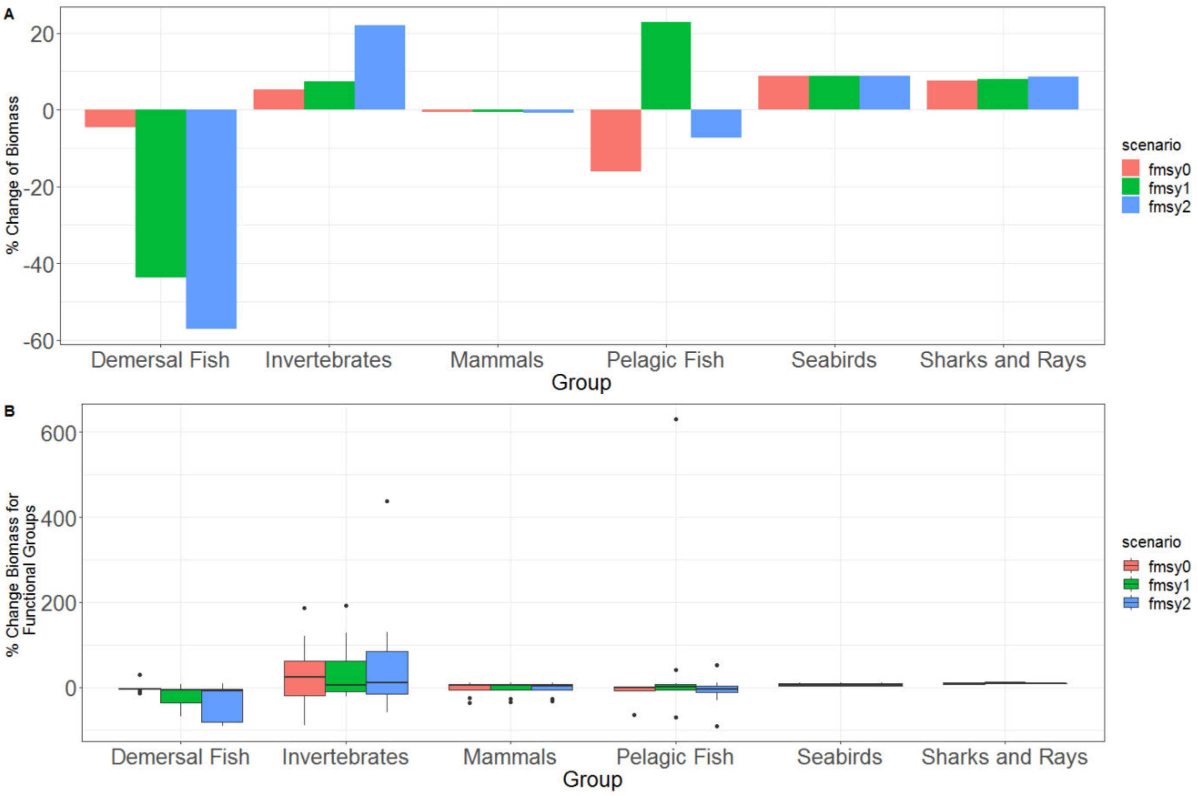

Figure 4-5: Changes in biomass of different species assemblages over 35 years under three different fishing scenarios in the NoBa Atlantis model. A) Shows the total change of biomass for each assemblage. B) Shows variation in the changes in biomass of different modelled functional groups within each species assemblage.

4.3.2 LPI and NNI trends

By displaying the indicators across the time period, we can better understand system changes over time (Figure 4-6). The LPI under the $F_{MSY}0$ shows the natural variation of the system with no fishing mortality applied; the LPI fluctuated over time, reflecting the trophic dynamics of the system, with an overall trend from 1981 to 2015 of -2% (95% CI -4.4%, -0.4%). When applying a fishing pressure of $F_{MSY}1$, the LPI shows an initial relatively steep decline in biodiversity followed by a more gentle decline from 1991 to 2008, followed by an increase in the final years, for an overall trend of -6.4% (95% CI -10.7%, -1.7%). These are in contrast to the scenario of $F_{MSY}2$ which shows a sharp decline in the LPI across the entire time period, producing a 26.1% (95% CI -30.4%, -21.8%) overall decline.

The regression analysis allowed us to generate 95% confidence intervals around the trend to test the extent to which each indicator was able to distinguish between the scenarios and better to understand the overall trend. A polynomial regression including scenario was significant for the LPI ($F(11, 303) = 305.6$, $p < 2.2e-16$). The LPI can clearly distinguish between the three scenarios, displaying significantly different trends for each fisheries scenario (Figure 4-7). This shows clearly that the LPI differentiates between the three scenarios and that fishing at $F_{MSY}2$ is most damaging to the ecosystem. The indicator picks up ecosystem changes very quickly; all three scenarios diverge within three years. The LPI indicates that the ecosystem stays relatively stable across the time period under no fishing, while fishing at $F_{MSY}1$ causes an initial decline in the ecosystem, which then stabilises. Fishing at $F_{MSY}2$ declines strongly and shows no sign of stabilisation, indicating that the ecosystem would decline further in time.

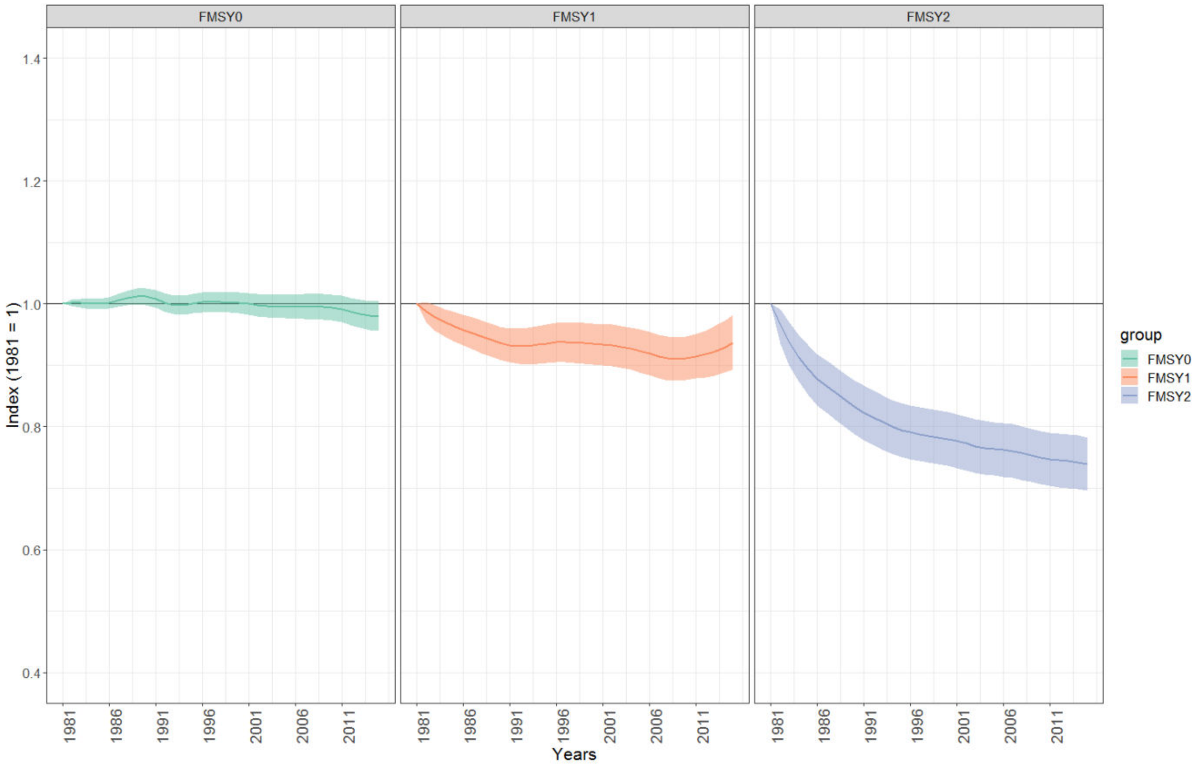

Figure 4-6: Responses of the modelled LPI to three different fishing pressures within the NoBa Model

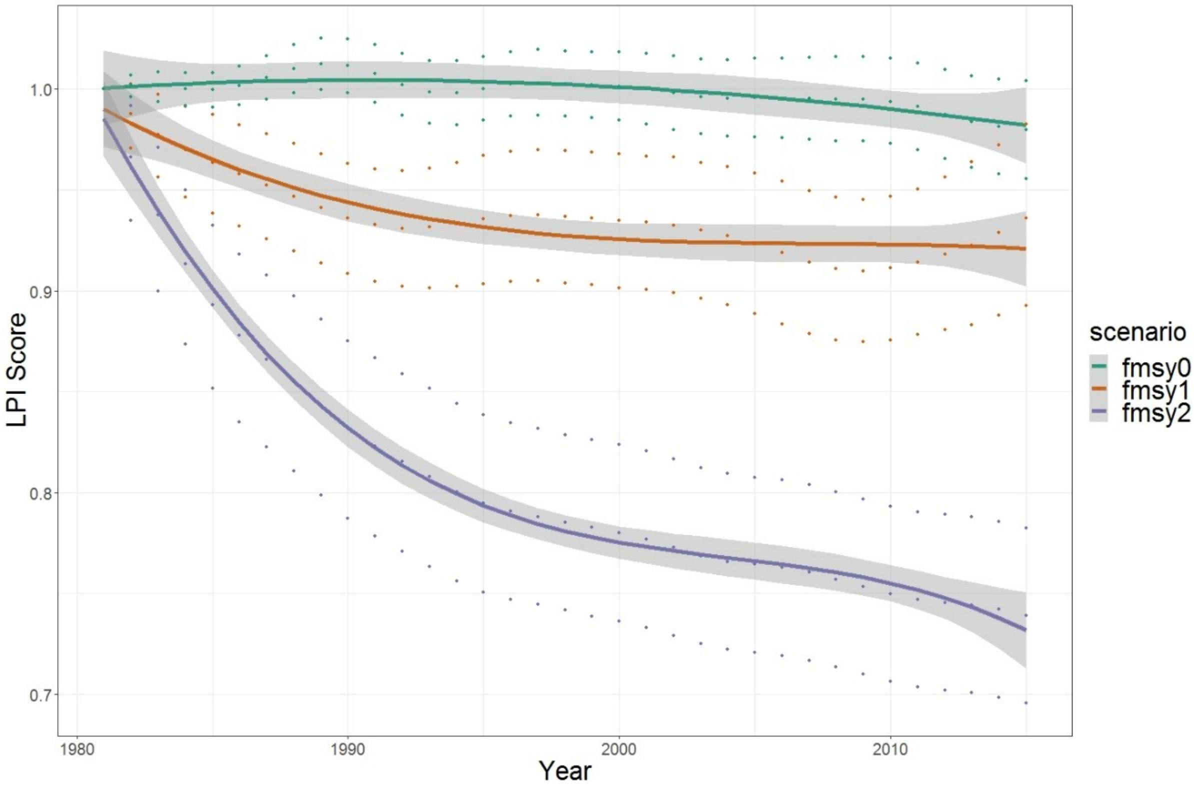

Figure 4-7: Regression analysis of LPI under three fishing scenarios. Grey shading denotes 95% confidence intervals and dots represent mean, upper and lower LPI scores.

The modelled NNI decreased overall by 13% under the $F_{MSY}2$ scenario and 5% under $F_{MSY}0$, although it increased by 0.7% under $F_{MSY}1$. While in the $F_{MSY}2$ scenario, the NNI showed a relatively consistent decline, under $F_{MSY}0$ and $F_{MSY}1$ it followed a general pattern of an initial increase, followed by a decrease (Figure 4-8). When breaking the NNI down by ecosystem type, the effects of fishing are clearer within the index. The Benthic NNI showed an increase of 1% under $F_{MSY}0$ and a decrease of 3% and 15% under $F_{MSY}1$ and $F_{MSY}2$ respectively. The Pelagic NNI decreased by 7% under $F_{MSY}0$ and 19% under $F_{MSY}2$, while it increased by 1% under $F_{MSY}1$. Confidence intervals are not available for the NNI as no uncertainty estimation is included. The variation in the pelagic NNI, which subsequently affects the overall NNI, is largely down to the inclusion of species with high natural variation such as zooplankton and mesopelagic fish. Invertebrate species such as zooplankton are not included within the LPI, and species such as mesopelagic fish carry more weight within the NNI than the LPI (which treats all species equally), due to the aim of weighting for ecological representation.

A multiple linear regression including scenario was significant for the NNI overall ($F(8, 96)= 21.52$, $p<2.2e-16$), the NNI pelagic ($F(11, 93)= 19.06$, $p<2.2e-16$) and the NNI benthic ($F(8, 96)= 608.4$, $p<2.2e-16$) (Figure 4-9). The NNI indicates that the $F_{MSY}2$ scenario is significantly more damaging for biodiversity for the overall score, the benthic score and pelagic score, however it was unable to distinguish between $F_{MSY}0$ and $F_{MSY}1$ for the pelagic and overall scores; even with a narrowing of the confidence intervals, the trend is unlikely to be distinguishable between the scenarios given the overlapping nature of the trend. The NNI appears to be much less responsive than the LPI, taking eight years to show differentiation between the three scenarios across all three formats of the NNI.

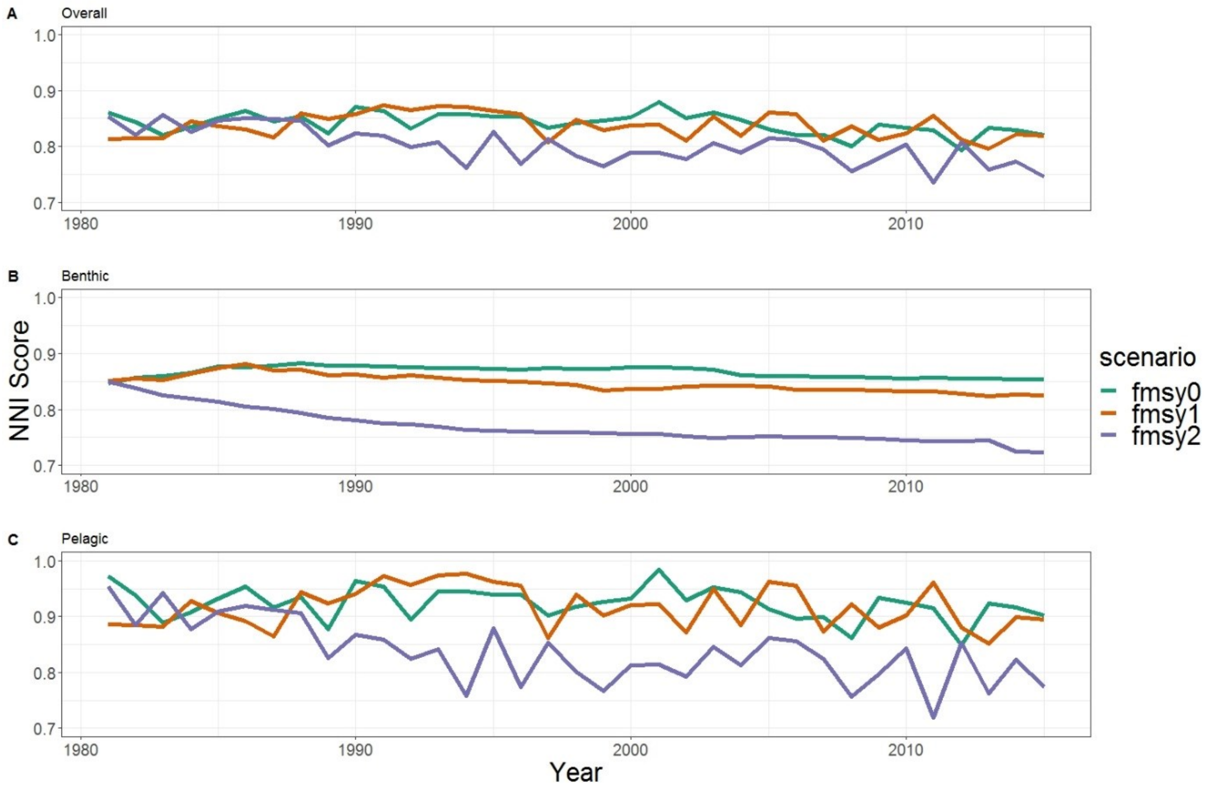

Figure 4-8: Responses of the modelled NNI to three different fishing pressures within the NoBa Model

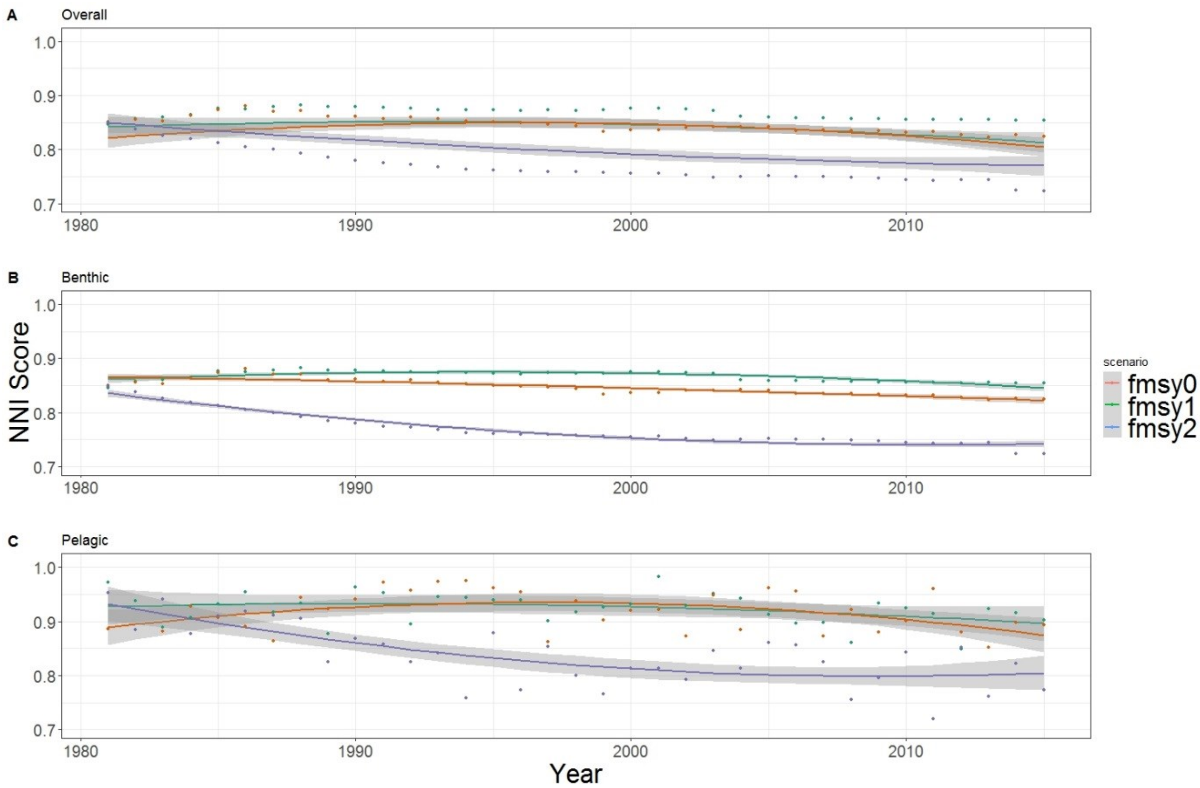

Figure 4-9: Regression analysis of the NNI under three fishing scenarios. Grey shading denotes 95% confidence intervals and dots represent NNI scores per year.

4.4 Discussion

Biodiversity is dynamic and constantly changing with and without human intervention. This makes the design, selection and use of indicators of the effectiveness of biodiversity conservation policy challenging, particularly at large scales. Nonetheless, there has been little evaluation of the ability of indicators to track underlying trends of interest, and to disentangle natural variation from human-induced changes in biodiversity. The LPI takes a best-available data approach, amalgamating as many vertebrate data sources as possible (which meet selection criteria), and applying a relatively simple weighting scheme. Alternatively, the NNI takes an ecosystem approach, targeting species that are most representative and applying a complex weighting scheme, based on species function and trophic level.

Overall, the indicators respond differently; the LPI is sensitive to change and quickly changes under the management scenarios. The NNI, however, takes longer to respond and finds no difference for biodiversity between the no fishing and sustainable fishing scenarios. In the example, both indicators signalled that the most heavily impacting management strategy was F_{MSY}2. This direct removal of biomass from the system caused a decline in biodiversity even once fishing effort had stabilised, showing the ongoing effects on the system of large human-induced mortality. However, the difference between F_{MSY}1 fishing and no fishing was not clearly pronounced. While fishing at this level caused a substantial decline in target species (Figure 4-4), the ecosystem in general at the guild level remained stable, with certain species benefitting from this level of fishing effort and others declining. The resultant trend in the LPI pointed to a decline in biodiversity between F_{MSY}1 and no fishing, however this was not the case for the NNI, which signalled no difference between no fishing and F_{MSY}1 for the overall score and pelagic score. While the LPI represented all vertebrate functional groups within the model, the NNI represented fewer functional groups overall, but contained nine invertebrate functional groups including phytoplankton, zooplankton, sponges and coral. As an assemblage, invertebrate species increased their biomass under all scenarios, showing underlying ecological conditions were favourable during the modelled period; in particular increasing water temperature from climate change has allowed for range expansion and increased productivity of these groups (Dalpadado et al. 2014). The LPI did not account for invertebrates and therefore was more sensitive to the declining target species and not accounting for externally-driven increases in abundance of other species. When separating out by ecosystem, however, there was a clear differentiation between the management scenarios for the Benthic NNI, which contained many of the exploited fish species. The Pelagic NNI contained many of the species with high inter-annual variability such as zooplankton, phytoplankton and small forage fish, which clouded indicator signal.

4.4.1 Linking models and indicators

Global biodiversity models do not yet contain sufficient information to enable species metrics to be modelled at the global scale (Hill et al. 2016), unlike global climate models, for example (IPCC 2014). As such, validating indicators at regional or national scales should be a priority for determining which should be scaled up or down for tracking global biodiversity targets. This will require different models being deployed in a number of regions and exploring a variety of indicators. Marine modelling is becoming increasingly advanced, with multi-modelling studies providing interesting insights into the difference between different models and regions (Olsen et al. 2018; Tittensor et al. 2018). Terrestrial modelling has tended to lack the mechanistic elements that are regularly explored in marine models, although new developments in this area show this may be improving (Rangel et al. 2018). If wide testing and validation of indicators can be achieved, then consensus on which indicators to use for which purpose will be more likely. Indicators can then act as a tool to guide and align species data collection, seeing use at multiple scales from local to global (Kissling et al. 2017).

Importantly, this co-development of indicators and models will then allow biodiversity indicators to be used for decision assistance. By projecting forward different scenarios, expected biodiversity change as a result of policy interventions can be estimated in an agreed, transparent and understandable manner, making explicit how policy decisions could impact biodiversity. By understanding the underpinnings of the indicator, as I have demonstrated here, we can predict the way in which different components of biodiversity might change, provoking discussions of how we would like biodiversity to look in the future. In the example, the NNI detect little change between no fishing and $F_{MSY}1$, while the structure of the ecosystem is being altered. By utilising a modelling approach, we can obtain a representation of the ecosystem, understand these changes and observe how the indicator responds.

In the real world, as with our example, there are inevitably many trade-offs between biodiversity and human activity but also within biodiversity itself. Overfishing has led to the reduction in marine biodiversity in many parts of the world, with invertebrate species replacing demersal fish in the ecosystem. This has led to new productive fisheries with high commercial value, such as for lobster, where a shift back to traditional fisheries might result in socio-economic loss (Howarth et al. 2014). Deciding whether society prefers mixed finfish ecosystems or invertebrate-dominated ones inevitably incurs trade-offs, but by better understanding how indicators used to make these trade-offs will pick up trends, society can be more explicit about what the future can look like and the paths we should take to get there.

The complexities of species-based mechanistic modelling, particularly at the global scale, and the general disconnection of this modelling to political targets and indicators has perhaps slowed progress in both using indicators to drive policy and monitor policy effectiveness. This work was possible because of the years of work developing the Atlantis modelling framework generally and the NoBa model specifically (Fulton et al. 2011a; Hansen et al. 2019). While such end-to-end models are not universal, they have significant global coverage and are constantly being developed, signalling great scope for advancing indicator testing with scenario analysis (Fulton 2010). Models should not, however, simply be believed and the benefits of this integration are both ways. The values that are incorporated within biodiversity indicators are not always reflected in the models (e.g. the LPI looks to measure all vertebrate life, no matter its ecological function, whereas NoBa is intending to be a representative sample of an ecosystem); by modellers and conservation scientists working together, models can be adapted to better parametrise certain aspects of conservation value (Wood et al. 2018). For example, while the model was primarily produced for exploring fisheries examples, with further work and assistance from conservationists, the number and quality of functional groups for sea birds could be greatly improved.

4.5 Conclusions

George Box's adage of "all models are wrong, but some are useful" is apt for biodiversity indicators. While many individuals and groups continue to generate new indicators, often striving for the impossible task of finding one that is "right", there is a lack of evidence to suggest which indicators are or are not practically useful. To have practical value, indicators must be designed carefully and clearly linked to what they are measuring, and the policy question they are meant to answer (Mcowen et al. 2016). While the CBD 2010 Framework committed to "immediate testing" of indicators, with a handful of exceptions this has largely not been enacted. This is potentially due to the complexities of doing so and the disconnection of conservation indicator scientists with ecosystem modellers, but may mean useful indicators are being side-lined or bad indicators are being adopted. Here I have provided a relatively straightforward example of how existing models and indicators can be brought together and adapted to help understand the drivers of ecological systems and indicator performance at the national scale. Wider adoption and application of such techniques would allow for better consensus on biodiversity indicators, can guide data collection to fill important data gaps, and assist in the production of science-based targets for biodiversity at a range of scales.

5 ASSESSING BIODIVERSITY LOSS WITH FISHERIES AND CONSERVATION INDICATORS

5.1 Introduction

Halting and reversing the loss of biodiversity is of critical importance to maintaining healthy natural systems where both nature and people are able to thrive (Watson et al. 2018). As such the principle of biodiversity restoration has been embedded as a core component in international agreements such as the Convention on Biological Diversity (CBD) and the Sustainable Development Goals (SDGs), particularly SDG 14 and SDG 15. While there is widespread agreement that biodiversity loss must be halted, how to best achieve this is less clear, as efforts to conserve biodiversity comprise a patchwork of international goals, national-level plans, and local intervention which overall are failing (Ripple et al. 2017). In marine systems, fisheries managers and conservationists play a key role in maintaining and preserving biodiversity, but often remain disconnected, operating independently of each other and having different histories, values and epistemologies (Salomon et al. 2011). As such these two sectors have typically had different objectives, tracking marine systems with different indicators and implementing different management measures, which are often not reconciled (Davies & Baum 2012). Being able to compare how management interventions impact both fisheries and conservation objectives is crucial for halting biodiversity loss while maintaining acceptable ecosystem services for people (Friedman et al. 2018). Making robust predictions about the future impact of human interventions is particularly important as the interactions between climate change and other anthropogenic impacts on biodiversity, such as fishing, remain unclear (Engelhard et al. 2014; Segan et al. 2016).

Norway is a maritime nation with a long history of fishing, which has been closely managed to ensure the future prosperity of fisheries resources (Grønnevet 2016). Norway has implemented knowledge-based fisheries management, based on the principle of sustainable harvesting, which has been the main management process for the preservation, restoration and continuation of marine ecosystems in Norway's waters. Sustainably managing fisheries, including wider ecosystem considerations, is seen as a key tool for Norway to meet its international obligations for species under the CBD (Norwegian Government 2018). This approach seems to be favoured over marine protected areas (MPAs) as Norway has designated just 1% of its territorial waters as MPAs, far below the 10% required under the CBD (Protected Planet 2019). In such a case it is important to assess how fisheries management approaches contribute to biodiversity performance and how consistently biodiversity indicators perform between sectors. Here we use the Nordic and Barents Sea (NoBa) Atlantis Model to explore the possibility of halting and reversing biodiversity loss into the future by simulating three different management approaches, which include climate change projections at RCP4.5, to 2068. The three different management approaches are Global Sustainability (fishing at $F_{MSY}1$), Precautionary Fishing (fishing at $F_{MSY}0.6$) and Strict Conservation (no fishing). I generate a range of indicators from across fisheries management and conservation to analyse progress towards halting and ultimately restoring biodiversity loss at three key timepoints; 2030 (the timeline for the SDGs), 2050 (the overall vision for the CBD) and 2068 (the end of model run). I likewise generate economic indicators to consider the trade-offs between biodiversity and socio-economics inherent in these management interventions. While also assessing management performance of each scenario we can also compare how different biodiversity indicators respond. By taking a modelling approach, we can compare changes in the indicators to what is happening in the 'virtual truth' of the model, which helps greatly with the interpretation of indicators which is so difficult in the real world (Nicholson et al. 2012).

5.1.1 Study Area

The Barents Sea is an open sub-Arctic shelf ecosystem situated north of Norway and north-west of Russia that covers an area of 1.6 million km^2 with an average depth of 230m (Dalpadado et al. 2014). The Barents Sea connects to the Norwegian Sea to the west and the Arctic Ocean to the north. The Norwegian Sea has a surface area of about 1.1 million km^2 with an average depth of about 1800m, resulting in a total volume of about 2 million km^3 (Loeng & Drinkwater 2007). Although they are designated as separate Large Marine Ecosystems (LMEs), the Barents and Norwegian Seas are highly interconnected. The physical conditions in both seas depend to a large extent on the inflow of Atlantic Water and share similar atmospheric driving forces. Herring (*Clupea harengus*) is a both culturally and economically important fish stock, which spawn along the Norwegian coast, using the Barents Sea as a nursery area, and feed in the Norwegian Sea as adults. Similarly, one of the world's largest stocks of cod (*Gadus morhua*) also spawns along the Norwegian coast, and their larvae drift with the currents into the Barents Sea, where they remain through their adult life. The seas also contain huge concentrations of seabirds and a diverse assemblage of marine mammals.

The area is considered to be exceptionally well managed by fisheries standards and continues to provide large catches of cod and other species well within safe and sustainable biological parameters as assessed in 2018 (ICES 2018). This is largely down to good management and long-term cooperation with neighbouring Russia (Grønnevet 2016). Despite fisheries success, some seabird populations have been declining rapidly (Fauchald et al. 2015) and sea ice loss in the Barents Sea has accelerated, with potential impacts on many species, but particularly marine mammals (Laidre et al. 2015).

5.1.2 Scenarios

The three scenarios were selected as a range of potentially plausible management interventions when thinking about fisheries and biodiversity. The first scenario, Global Sustainability, undertook fishing at $F_{MSY}1$. $F_{MSY}1$ is the maximum rate of fishing mortality that will eventually result in a population size of B_{MSY}. B_{MSY} is the biomass that enables a fish stock to deliver the maximum sustainable yield and is often used as a reference point for managing fisheries. Although the concept of Maximum Sustainable Yield (MSY) is somewhat controversial (Rindorf et al. 2017), MSY is a key part of the United Nations Convention on the Law of the Sea (UNCLOS) :

"…State(s) must set an allowable catch, based on scientific information, which is designed to maintain or restore species to levels supporting a maximum sustainable yield (MSY)."

(United Nations General Assembly 1982)

As such it has gained global support as a way of achieving sustainability in fisheries. The second scenario, Precautionary Fishing, includes fishing at $F_{MSY}0.6$, which might be considered a rather low rate of fishing. Fishing below MSY in theory allows fish stocks to recover if overfished or adds a precautionary buffer so that they do not subsequently become overfished. The third scenario is Strict Conservation, which ended all fishing in the model. Such a scenario is interesting as it would be considered to be pro-biodiversity, but with significant socio-economic and cultural impacts.

5.2 Methods

5.2.1 Model Overview

An updated version of the NoBa model, as described in Chapter 4, was parameterised with historical fishing data from 1980-2017 and run forward under three different management intervention scenarios from 2017-2068. This meant that baseline conditions at the implementation of management interventions were intended to be representative of a situation similar to reality, unlike Chapter 4, where the starting conditions represented an unfished system. Although end-to-end models are never finished, the NoBa model has undergone sensitivity to analysis to explore parameter uncertainties, which is largely uncommon for end-to-end models (Olsen et al. 2016; Hansen et al. 2019). All scenarios included historic temperature and salinity data, projected forward to 2068 under Representative Concentration Pathway (RCP) 4.5 from the Regional Ocean Modelling System (ROMS; Shchepetkin & McWilliams 2005) covering the Northeast Atlantic (Skogen et al. 2007). This represents a moderate climate change scenario, with peak emissions around 2040, and then declining. Fishing mortality was included for 11 commercial species within the model, which was ended for the Strict Conservation (no fishing) scenario. The other two scenarios applied fishing mortality at $F_{MSY}1$ (Global Sustainability), and $F_{MSY}0.6$ (Precautionary Fishing) for each of the fished species. It is not possible to calculate MSY for capelin (*Mallotus villosus*) as it is a short-lived species which dies after spawning. For this species, mortality was included by taking the average annual catch from 2007-2017 to represent $F_{MSY}1$ and multiplying by 0.6 for $F_{MSY}0.6$, on the assumption that the species is currently fully but sustainably exploited.

5.2.2 Indicator Overview

I generated three types of indicators used to measure marine systems to represent how different sectors might monitor the effect of management on biodiversity trends. These included conservation indicators, fisheries ecosystem indicators and IndiSeas indicators, which are a general set of indicators used for comparing biodiversity impacts across marine systems. I also generated a selection of fisheries/economic indicators used regularly within fisheries management to monitor fisheries outputs.

For the conservation indicators, I generated the Living Planet Index (LPI) and Norway Nature Index (NNI) similarly to the approach set out in Chapter 4. The LPI represents a typical abundance-based indicator, which can be translated for use at the national scale and easily reported back to the global scale. The NNI was developed by Norway for use at the national scale to track its biodiversity performance in light of international commitments like the CBD. The LPI requires a time-based reference point which was set as 2015, as this is when the SDGs were established and could be considered a suitable timepoint to measure against. The NNI typically requires a reference point for each species in what would be considered its ‘undisturbed’ state. From a modelled perspective this would be difficult to estimate for these scenarios as they are parameterised with historical fishing data. As such, to keep in line with other indicators, I established the reference point for each species as the average biomass in 2010-2014, thus setting 2015 as the baseline reference against which future biodiversity loss could be measured. I also generated a third conservation indicator measuring the state of iconic species. The Ocean Health Index contains iconic species as a sub-goal to the Sense of Place Goal (Halpern et al. 2012), which is calculated from IUCN Red List status. I calculated the relative abundance of iconic species, based on Fulton et al. (2018). This indicator included typical species of conservation concern such as vulnerable (slow growing) species – in this case it included all cetaceans, pinnipeds and seabirds. Higher scores for this indicator convey that the system structure has not been distorted by the loss of these vulnerable and culturally important species.

I also generated ecosystem indicators used by the IndiSeas project to reflect the ecological and biodiversity status of marine ecosystems (Shin & Shannon 2010; Shin et al. 2010a, 2010b). A key aim of this project was to provide a set of synthetic indicators, which could be used comparatively across different ecosystems of the world (Shin & Shannon 2010). The initial suite of IndiSeas indicators was supplemented with a list of empirically-based candidate biodiversity indicators initially established based on ecological significance (Coll et al. 2016). IndiSeas indicators have been subject to testing and analysis to support their selection (Link et al. 2010; Shin et al. 2018), but reference point selection has proved a challenge as they have been primarily designed for comparison between different systems (Shin et al. 2010a). In this case, I use a time-based reference point of 2015 for comparing progress for each scenario. Higher values of the IndiSeas indicators correspond to good ecological status of ecosystems, and improving scores corresponds to positive trends in indicators. They contain both ecosystem-based and catch-based indicators. A full list and description of IndiSeas indicators can be found in Table 5-1.

Table 5-1: Description of IndiSeas indicators generated from the NoBa Model

Indicators	Descriptor	Notations
Total biomass of surveyed species	*biomass*	*B surveyed* (tons)
1/(landings /biomass)	*inverse fishing pressure*	B/Y retained species
Trophic Level landings	*trophic level landings*	$TL_{land} = \frac{\sum_s TL_s Y_s}{\sum_s Y_s}$
Proportion of predatory fish	*% predators*	prop predatory fish= B predatory fish/B surveyed
Mean life span	*life span*	$\frac{\sum_s (age_{max} B_s)}{\sum_s B_s} (year)$
TL Community	*Trophic level Community*	$TL_{comm} = \frac{\sum_s TL_s B_s}{\sum_s B_s}$
s: species, N: abundance, B: biomass, Y: catch, TL: trophic level Surveyed species are species sampled by researchers during routine surveys (as opposed to species sampled in catches by fishing vessels), and include species of demersal and pelagic fish (bony and cartilaginous, small and large), as well as commercially important invertebrates (squids, crabs, shrimps) (IndiSeas 2019).		

Ecological and biodiversity indicators for fisheries have been widely discussed, tested extensively and used in a variety of studies (Fulton et al. 2005). In addition to the indicators listed above I also chose a set of six indicators which are widely used to track status and trends of ecosystems and have been used in various other modelling studies analysing ecosystem responses to fishing pressure, which can be found in

Table 5-2 (Fay et al. 2013; Olsen et al. 2018). Good ecological status of ecosystems is characterized by high values of these indicators, and improving scores corresponds to positive trends in indicators.

Table 5-2: Description of fisheries ecosystem indicators generated from the NoBa Model

Indicators	Description	Notations
Pel bio/PP	Ratio of pelagic biomass to primary production	B pelagic / B primary production
Bio/PP	Ratio of total biomass to primary production	B / B primary production
Dem/Pel	Ratio of demersal to pelagic fish biomass	B demersal fish / B pelagic fish
Dem bio/PP	Ratio of demersal biomass to primary production	B demersal / B primary production
Prop Pel	Proportion of total biomass that is made up of pelagic species	B pelagic / B
PropPred	Proportion of total biomass that is comprised of predatory species	B predators / B

To explore socio-economic changes under the scenarios I used a series of indicators which measure fisheries and economic properties and can be used as a proxy for the socioeconomic effects of management decisions (Olsen et al. 2018). These indicators can be found in

Table 5-3.

Table 5-3: Description of fisheries indicators generated from the NoBa Model

Indicators	Description	Notations
Total Catch	Summed catch of all species	Y
Pelagic Catch	Summed catch of pelagic species	Y pelagic
Demersal Catch	Summed catch of demersal species	Y demersal
Fish Catch	Summed catch of fish species	Y fish
Fish Exploitation Rate	Total catch/Fish Biomass	Y/B fish
Target Species Exploitation Rate	Total Catch/Target Species Biomass	Y/B target species

All ecosystem indicators were measured annually through time and then compared to their 2015 baselines (2010-2014 average) in 2030 (measured as average of 2025-2029 inclusive), 2050 (measured as average of 2045-2049 inclusive) and 2068 (measured as average of 2063-2067 inclusive) to see to what extent that they reported on biodiversity loss being halted. This enabled me to pinpoint and interpret similarities and differences across scenarios and indicator groups. In the case of halting biodiversity loss, there is no consensus on indicator reference points. Comparing indicators against time-based reference points is one of the most common ways of measuring progress against policy goals, which often lack science-based reference points (McQuatters-Gollop et al. 2019). As such I assume a time-based reference point of 2015 and any indicators that fall below this reference point indicate that biodiversity loss has not been halted. By taking such an approach, I can clearly communicate results and compare between indicators to a broad audience, in keeping with the purpose of indicators (ten Brink 2006).

5.3 Results

5.3.1 Ecosystem Dynamics

The three different management scenarios altered the ecosystem in different ways, particularly at the functional group level (Figure 5-1). The most significant changes were to commercially caught species, such as the Greenland halibut (*Reinhardtius hippoglossoides*), which directly responded to the changes in fishing pressure as their mortality was increased or decreased depending on the scenario and their current fishing pressure. Other species such as the Long rough dab, (*Hippoglossoides platessoides*), although not under any direct fishing mortality, had its biomass altered significantly by the management scenarios. This is because of its close relationship to those species that are caught commercially; as it is outcompeted by them as their numbers increase from reduced fishing pressure under Strict Conservation. Other species, such as the Killer whale (*Orcinus orca*), showed less response to the management scenarios but still displayed strong underlying trends over time as a result of the climate change built into the model.

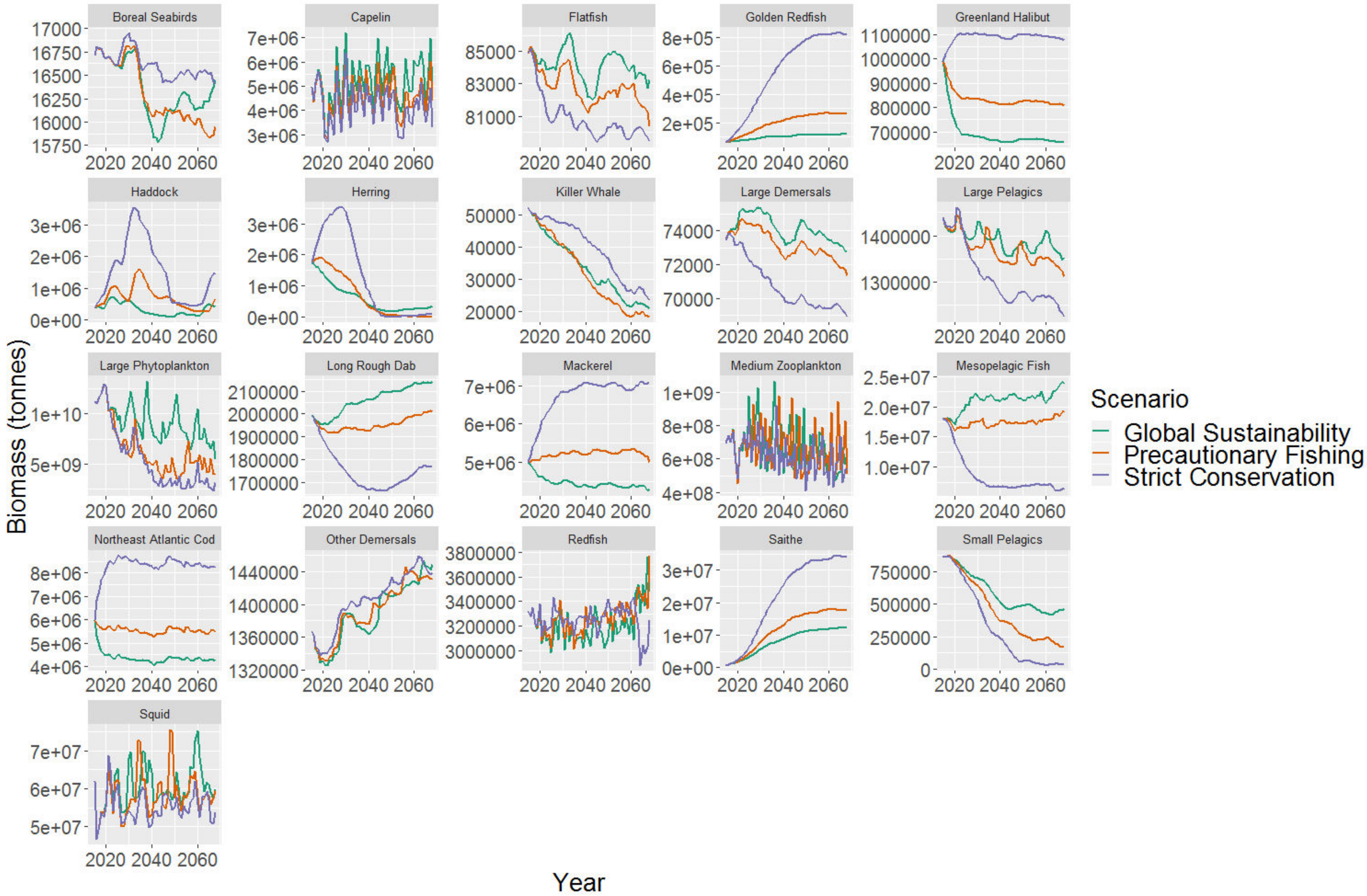

Figure 5-1: Responses of selected functional groups to three different fishing scenarios in NoBa model

Underlying trends are clearer when looking at the guild level (Figure 5-2). Many guilds show little differentiation between the management scenarios, showing there are strong ecosystem dynamics underpinning and stabilising the relationships in the system, driven by historical exploitation, interactions between species and also the added effect of climate change. For instance, filter feeders such as coral and sponges increased strongly under all scenarios; more than doubling their biomass by 2068, probably due to increased productivity under climate change. Demersal fish show one of the greatest changes between scenarios, increasing 22.6% by 2030 under Strict Conservation, compared to an 8.2% decrease under Global Sustainability. However, this pattern was not uniform across guilds; primary production showed a large decreasing trend under Strict Conservation, reducing by 74% by 2068, compared to a 3.6% increase under Global Sustainability.

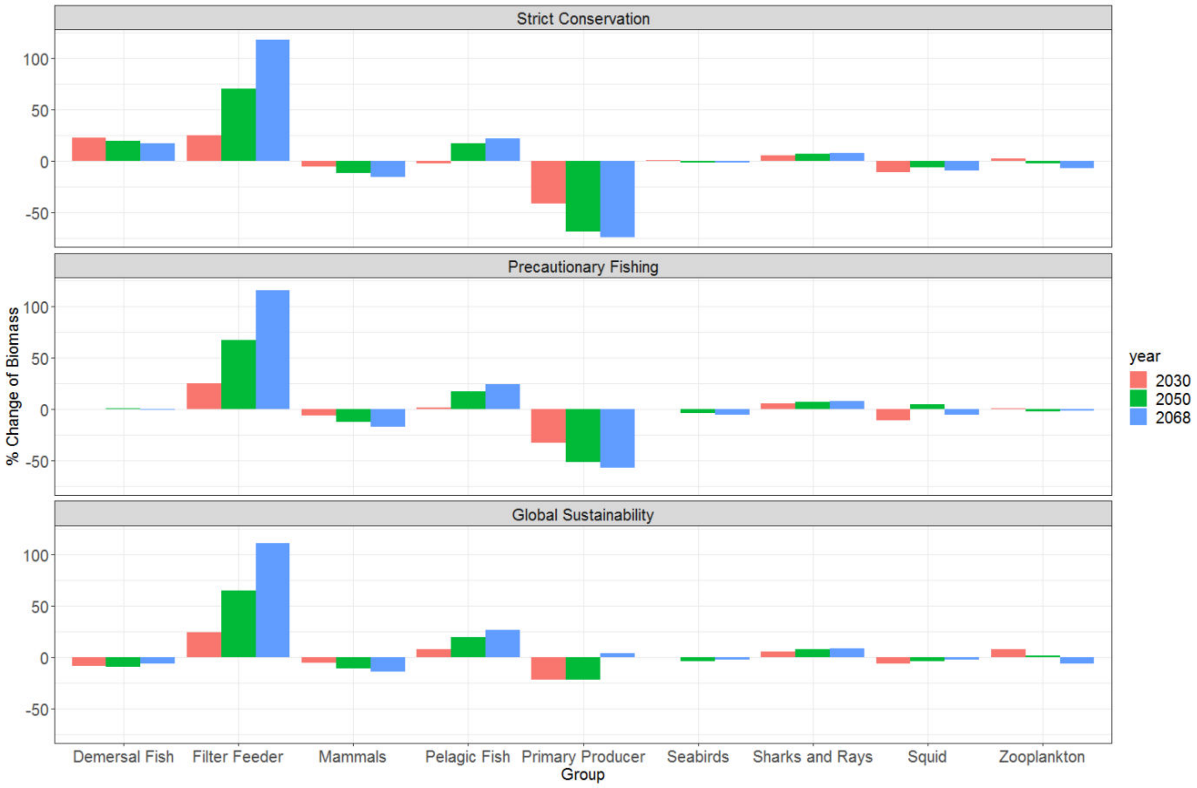

Figure 5-2: Guild level biomass responses at three time points across three fishing scenarios from the NoBa model

5.3.2 Indicators

A summary of whether each indicator reports that biodiversity loss has been halted at key timepoints is shown in Table 5-4. Exact values and full timeseries can be found in Appendix 4. When comparing against 2015 baselines, by 2030, one indicator reported that biodiversity loss had not been halted under Strict Conservation, four under Precautionary Fishing, and six under Global Sustainability. Every indicator except total surveyed biomass had Strict Conservation as the best performing scenario to 2030. Indeed, when exploring the total surveyed biomass indicator, and removing zooplankton (Figure 5-3), the signal became much clearer and Strict Conservation became the best performing scenario across all timepoints. Zooplankton are regularly surveyed in Norwegian waters as they are being considered for a potential new fishery (Samuelsen et al. 2009) but have high interannual variability and therefore can act as noise in indicators. Only iconic species abundance fell below the baseline level by 2030 for Strict Conservation; this indicator would be unlikely to see any immediate response in change to fishing pressure because of time lags with long-lived species. However, responses from the other indicators shows that the ecosystem overall responded quickly and positively to such a dramatic change in policy, with the indicators largely in agreement that this was positive for biodiversity. For the two fishing scenarios to 2030, the IndiSeas indicators generally pointed towards halting biodiversity loss being achieved, whereas the fisheries ecosystem indicators displayed mixed messages, with large gains in the primary productivity indicators and losses in the others (PropPel, PropPred, DemPel).

From the conservation indicators, the LPI and NNI perform relatively similarly, as both indicators respond positively under Strict Conservation and Precautionary Fishing to 2030, showing that they respond in a timely and reliable manner to these management changes. A positive response under Strict Conservation especially would be expected, as fishing mortality is halted. The LPI seems to have greater sensitivity, with a 10% increase under Strict Conservation, compared to 3.75% for the NNI, most likely due to the aggregation methods constraining the NNI’s sensitivity (see Chapter 4).

Table 5-4: Summary of whether indicators show biodiversity loss was halted at three time points across three fishing scenarios. Assessment was made by comparing each time point to the baseline level in 2015. (Y= Yes, N= No).

Scenario		Global Sustainability			Precautionary Fishing			Strict Conservation		
Year		2030	2050	2068	2030	2050	2068	2030	2050	2068
Conservation	LPI	Y	N	N	Y	N	N	Y	N	N
	NNI	N	N	Y	Y	N	N	Y	N	N
	Iconic Abundance	N	N	N	N	N	N	N	N	N
IndiSeas	Total Biomass	Y	Y	N	Y	N	Y	Y	Y	N
	% Pred	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Mean Life Span	Y	N	N	Y	Y	Y	Y	Y	Y
	TL Community	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Inverse Fishing Pressure	N	Y	Y	N	Y	Y	NA	NA	NA
	TL Landings	Y	Y	N	Y	Y	N	NA	NA	NA
Fisheries Ecosystem	PelBioPP	Y	Y	Y	Y	Y	Y	Y	Y	Y
	BioPP	Y	Y	N	Y	Y	Y	Y	Y	Y
	DemPel	N	N	N	N	N	N	Y	N	N
	DemPP	Y	Y	N	Y	Y	Y	Y	Y	Y
	PropPel	N	N	N	N	N	N	Y	Y	Y
	Prop Pred	N	Y	N	Y	Y	Y	Y	Y	Y

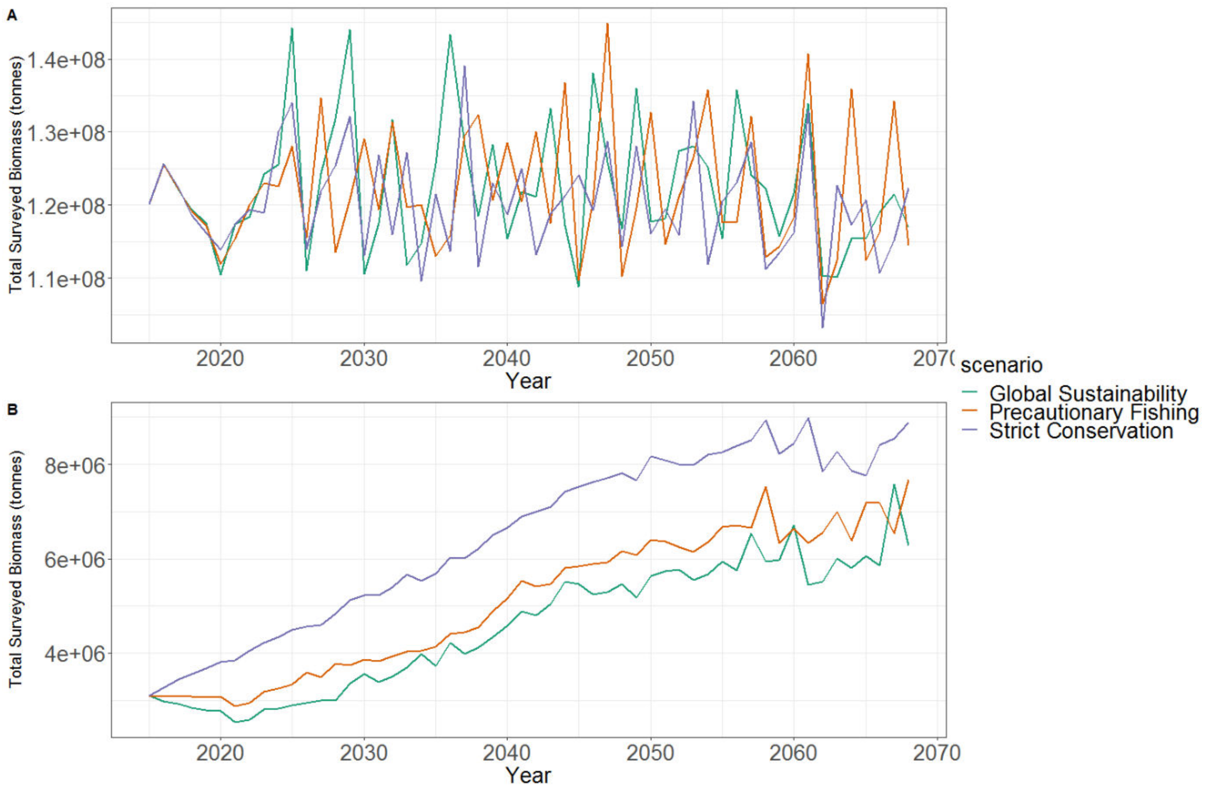

Figure 5-3: A) Total surveyed biomass across the model run for three different fishing scenarios B) Total surveyed biomass excluding three species of zooplankton

When looking to 2050, the indicators start to diverge more. Under Strict Conservation, four indicators point to biodiversity loss not being halted at 2050, compared to six each for the fishing scenarios. While Strict Conservation is still the best performing scenario for almost all of the IndiSeas and fisheries ecosystem indicators, it is not the case for the conservation indicators. Strict Conservation becomes the worst performing scenario for the NNI in 2050 and 2068, and middle performing for the LPI and iconic species abundance. The iconic species abundance indicator declined across all scenarios and timepoints showing that higher-trophic level groups such as cetaceans are much less impacted by the fishing scenarios and are exhibiting underlying declines in their biomass due to climate change. Many of the species included within the iconic species indicator were also included within the NNI and LPI and many declined consistently under all scenarios across the timeseries, further driving negative responses of these indicators. From the fisheries ecosystem indicators, those concerned with ratios of primary productivity (PelBioPP, BioPP, DemPP) showed gains under all scenarios, although with significantly higher gains under Strict Conservation. Primary productivity drops largely under the Strict Conservation management scenario (Figure 5-2), reflecting phytoplankton declines due to trophic interactions.

Conservation indicators show decline under all scenarios (except NNI under Global Sustainability) to 2068, whereas many of the fisheries ecosystem and IndiSeas indicators respond more as expected, with increasing scores under less fishing pressure. Five of six fisheries ecosystem indicators show large gains to 2068 under Strict Conservation. From a socio-economic perspective, fish catch indicators performed as expected (Figure 5-4). Catch initially dropped as the new management regimes were established but then in general steadily rose over time, driven by large catch increases of prawn (*Pandalus borealis*) and saithe (*Pollachius virens*), whose biomass increased largely over the model run in all scenarios, showing they have benefited from climate change. Trends in these two species drove pelagic catch (for saithe) and demersal (for prawns) upwards, but the increases were less apparent when looking at total fish catch which excluded prawns. Global Sustainability provided 17.1% more catch than Precautionary Fishing over the model run, which works out as a difference of over $38billion USD over the course of the model run or $1billion per year at 2015 prices. Catch data for all commercial species can be found in Appendix 4.

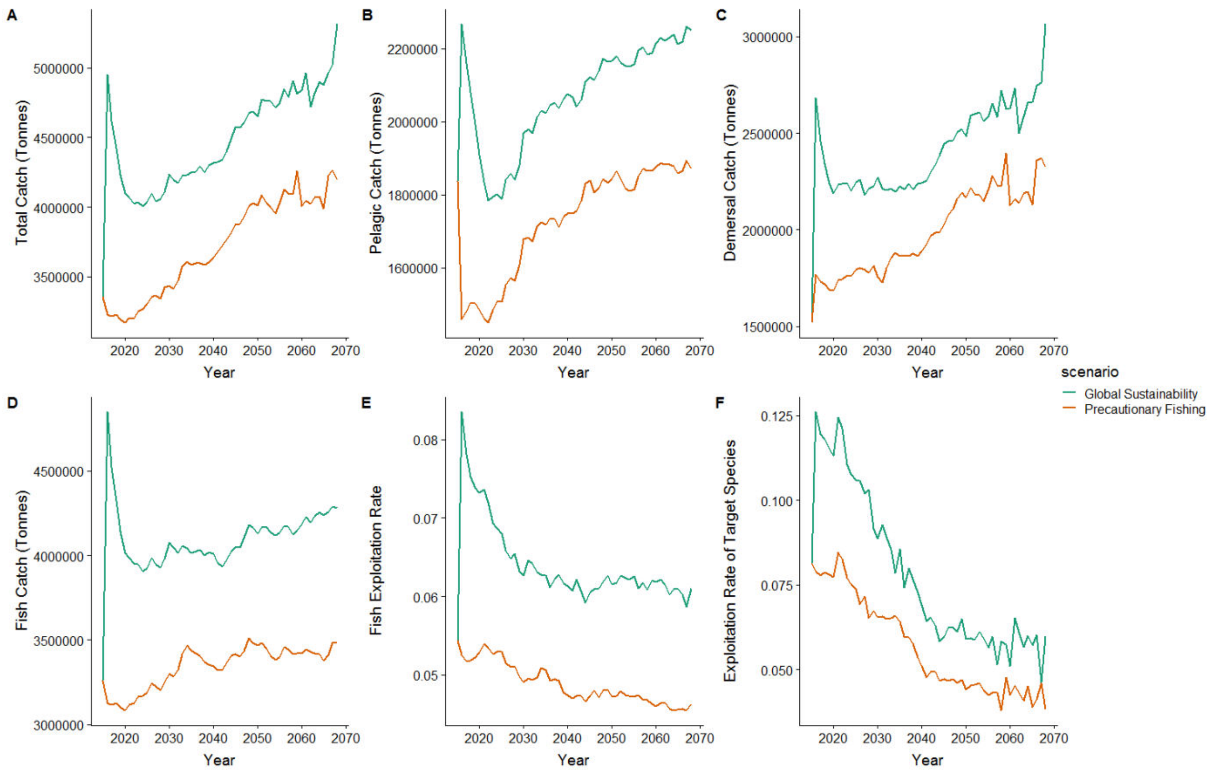

Figure 5-4: Fisheries indicator responses to two fishing scenarios generated from the NoBa model

5.4 Discussion

Here I applied three contrasting fisheries management approaches, within the region of the Nordic and Barents Seas over a period of 38 years. I did this with two aims; to see if they could halt and potentially reverse biodiversity loss in line with international commitments, and to see how consistent these predictions were between indicator types. I used a variety of biodiversity indicators from across different sectors to see how they compared and if different design philosophies underpinning them had an effect on how biodiversity is potentially communicated and managed.

5.4.1 Halting biodiversity loss

When analysing across the suite of indicators as a whole, they suggest that Strict Conservation offers the best outcomes for biodiversity to 2068 of all three scenarios; By 2068, nine out of 13 indicators point to biodiversity loss being halted under Strict Conservation, compared to ten and six out of 15 for Precautionary Fishing and Global Sustainability respectively. However, not all indicators reflect this, nor do any of the scenarios have consensus on whether or not international biodiversity commitments will be met within the NoBa area following these approaches alone.

All conservation indicators, bar the NNI under Global Sustainability, declined below the 2015 baseline by 2068 under all scenarios. The initial improvements in biodiversity shown by the conservation indicators at 2030 was thus not continued in the medium term (the timeframe of the strategic vision of the CBD). This pattern can be attributed to the stabilisation of fished species that initially showed positive responses from reduced fishing pressure, before density dependence limited further growth (e.g. North-east Atlantic Cod) (Andersen et al. 2017). Also, trophic effects were responsible for some fished species (e.g. herring) sharply declining following initial growth as they were heavily predated. Other species performed worst under Strict Conservation (e.g. small pelagics) as they were outcompeted by other rebounding species. The conservation indicators generally considered a much wider array of individual species, ranging from invertebrate species right through to seabirds and cetaceans, reflecting wider conservation concerns and interests. These were considered as individual species and then weighted differently depending on the indicator. From a conservation perspective, Global Sustainability led to the most stability over the time series and had the highest indicator values by 2068. This is somewhat surprising but shows the strength of underlying ecosystem dynamics, which shifted dramatically under management thought to favour biodiversity (Strict Conservation) and remained most stable over the scenario which was closest to current and historic management (Global Sustainability). In general, the fisheries ecosystem and IndiSeas indicators yielded improved scores with lower fishing effort, as would be expected. For some of the fisheries-based indicators (e.g. PelBioPP, % pred), all scenarios led to upwards trends over time, meaning in isolation, each scenario would be considered a positive outcome from a management perspective. However, by running different scenarios it can be seen that alternative options can offer different levels of change (e.g. % predators increased 284% under Strict Conservation compared to 78% under Global Sustainability).

5.4.2 Comparing fisheries and conservation biodiversity indicators

The conservation and fisheries-based indicators are constructed and used in different ways, both of which have advantages and disadvantages (Table 5-5). The development and testing of ecosystem indicators for fisheries has focussed on key attributes such as responsiveness and reliability for detecting change, primarily in fishing pressure, in order to try and advance the operationalisation of Ecosystem Based Fisheries Management (EBFM) (Tam et al. 2017). This has been a rather different approach to conservation indicators, which have been developed, usually in isolation, to reflect certain aspects of biodiversity that are widely cared about, such as extinction risk (Butchart et al. 2005), vertebrate abundance (Collen et al. 2009), intactness (Purvis et al. 2018) or ocean health (Halpern et al. 2012). The differences between these approaches is reflected in the results of this study, where conservation indicators paint a different picture to fisheries ecosystem and IndiSeas indicators.

Both IndiSeas and fisheries ecosystem indicators favour using several indicators as a suite with relatively straight forward construction methods, such as ratios, summations or proportions. This differs from the conservation indicators which tend to be aggregated indicators with more detailed methods and used in isolation. Conservation tends to be species-focussed and therefore attempts to collate and aggregate as much species level information as possible (Mace et al. 2008). The conservation indicators are predicated on an assumption that all species they include should increase in abundance in concert under favourable management, which makes their interpretation particularly challenging as species population dynamics are typically complex and change based on numerous factors, both natural and anthropogenic (Regan et al. 2002). I have demonstrated here that even when acting with biodiversity in mind (Strict Conservation), ecosystems do not respond in a linear fashion, and species interactions mean there are winners and losers (Hill et al. 2016). This is a challenge for aggregate indicators as they can mask these trade-offs. In the real world, depending on which species are captured by indicators, such as the LPI, which use all available data, there is potential for misleading or masking of signals (Nicholson et al. 2012). I have shown here, however, that by developing an LPI based on targeted data from the ecosystem, that it is a useful supplement to fisheries ecosystem indicators. However, indicators such as the NNI which seek ecological representation, may face problems in interpretation as they often use composite measures with high levels of weighting and perfect compensability, meaning negative changes in one part of the index can be offset by positives in the other which generally impacts sensitivity (Burgass et al. 2017). By integrating indicators within a modelling framework as done here, I have been able to compare responses in biodiversity indicators to the ecological changes within the model. This has helped interpret and thus validate the responses in conservation indicators, which is rarely considered (Costelloe et al. 2015).

The IndiSeas and fisheries ecosystem indicators approach differs as they measure different aspects of the ecosystem, often at the guild level or above. Results at higher taxonomic levels (Figure 5-2) are often more stable than when considered at the species level (Figure 5-1) (Olsen et al. 2018), and therefore these higher levels have been a major focus for fisheries indicator science (Fulton et al. 2005). However, as we see here, not all the indicators in a suite perform uniformly, requiring further interpretation. For instance, total biomass of surveyed species was highly variable and revealed little difference between the scenarios (Figure 5-3, A). However, when removing zooplankton, the differentiation between scenarios became much clearer (Figure 5-3, B), with Strict Conservation the best performing scenario. Likewise, there are two indicators that are concerned with measuring predators. % Predators from the IndiSeas project looks at the ratio of predatory fish compared with surveyed biomass, whereas the fisheries ecosystem indicators look at total predator biomass compared to overall biomass

(Prop Pred). Although both indicators increase over Precautionary Fishing and Strict Conservation, they exhibit different results under Global Sustainability, with Prop Pred declining by 11% by 2068 and % Pred increasing by 78%. In the case of the Prop Pred, the total system biomass is increasing at a greater rate than predators, thus causing the indicator to decline, despite predator numbers also increasing. Understanding how indicators' construction affects their performance is important, but often overlooked (Moriarty et al. 2018).

Table 5-5: Pros and cons of fisheries ecosystem and conservation biodiversity indicators

Indicator Type	Indicator attributes	Pros	Cons
Fisheries Ecosystem Indicators	Multiple indicators used as a suite	• Gain information on multiple attributes of the system	• Hard to synthesise/understand if indicator responses aren't the same • More complex communication
Conservation Indicators	Single indicators favoured	• Used more easily with political target setting • Easier to communicate with broad audiences	• Lacks full system understanding for decision making
Fisheries Ecosystem Indicators	Subject to rigorous simulation testing	• Indicators are sensitive and perform as expected • Able to project management approaches forward and communicate results using indicators	• Focus on high performance may exclude parts of the system that are societally valued e.g. certain species groups
Conservation Indicators	Built around aspects that are valued i.e. species	• Measure valued parts of the system	• Unclear how indicators perform in reality, which impacts their usefulness
Fisheries Ecosystem Indicators	Favour simple ratios of higher guilds	• Based on scientific understanding and easily interpreted	• Don't measure species changes
Conservation Indicators	Favour aggregations	• Can include large amounts of data	• Can hide trade-offs and dynamics
Fisheries Ecosystem Indicators	Data targeted for indicators	• Aids understanding of indicator	• Requires potentially expensive monitoring programmes • Narrower focus of species

Table 5-5: Pros and cons of fisheries ecosystem and conservation biodiversity indicators

Indicator Type	Indicator attributes	Pros	Cons
Conservation Indicators	Data used based on availability	• Can integrate and include broad data sets • Cheaper to develop	• Prone to geographic and taxonomic bias as well as data uncertainties, which can cloud signals

5.4.3 Using indicators

Norway has ambitious commitments to leave areas such as Svalbard 'virtually untouched', but at the same time to increase revenues from the blue economy, through tourism, fisheries, aquaculture and other sectors reliant on different aspects of biodiversity (Norwegian Government 2017). It is important that biodiversity indicators are meaningful and easily communicable because ultimately there are always trade-offs between fisheries, conservation and other aspects of social-ecological systems. In this example, lowering fishing pressure comes at a large economic cost and yet does not provide significant benefits for biodiversity according to conservation indicators. In such a case further action would be necessary outside of, but linked to, fisheries in order to meet biodiversity objectives. Cross-sectoral management is necessary to navigate trade-offs and achieve overall targets and this requires integrated planning and management. Norway has already initiated this process through its integrated management plans for the Barents and Norwegian Seas, but individual sectors continue to manage and report on their own objectives (Norwegian Ministry of Climate and Environment 2016). The NNI was created in order to bring different specialist groups together and assess biodiversity in a single metric, but as we have seen here, there is substantial value in reporting across a suite of indicators and factoring in multiple objectives.

One of the key challenges in using indicators is that they are often developed without specific management objectives in mind. As such, defining thresholds or trigger points for action has proved extremely difficult, even within fisheries science where indicators have been explored at length (Samhouri et al. 2010; Large et al. 2013). The LPI uses a time-based threshold and the NNI uses a population threshold, which allows comparison of how far we are from a desired state, but indicators do not help understand how to achieve the desired state or if it is even possible. Understanding the wider system, making predictions and undertaking adaptive management are all critical in actually achieving biodiversity outcomes. When juggling competing objectives, participatory methods are often important as value judgements are necessary (Punt et al. 2016). Integrating modelling scenarios and participation has proved successful for setting meaningful thresholds to trigger management intervention in marine systems (Addison et al. 2015). Without undertaking these processes, discussion and management will continue to be siloed and indicators remain poorly interpreted, as science alone cannot make such trade-offs.

5.5 Conclusions

Both fisheries and conservation are concerned with the preservation of biodiversity. In Norway, fisheries are the dominant sector for managing biodiversity, but the types of indicators used in this sector are not necessarily well aligned with conservation priorities. Wider system changes of conservation concern were not detected by any of the fisheries ecosystem indicators; in essence because they have been designed to be responsive to fishing pressure. This has advantages in that many of those indicators were extremely sensitive to change, but there are questions about whether they measure ecosystem components which are of particular cultural or economic (outside fisheries) importance and reflected more in the conservation indicators (Hobday et al. 2015). Indeed, many of the higher trophic level species are purposefully left out of these indicators, despite being of high conservation concern and/or interacting with fisheries (Davies & Baum 2012).

Indicators are required, not only to detect changes because of fishing, but all changes in marine ecosystems. This increases the complexity of indicators, not only in design but in interpretation. Conservation indicators tend to be aggregated metrics which appear simple and intuitive on the surface, but have been criticised for not acknowledging their uncertainties (Collen & Nicholson 2014). Fisheries indicator testing and use is highly advanced and the methods used could be of great benefit to wider conservation initiatives, particularly at large scales. However, it is important to ensure integrative participatory decision-making to understand preferences for marine systems. This will help further guide what should be measured and what other conservation actions are required. Equally, conservationists could benefit from awareness of recent advances in marine fisheries science and the complexities of ecosystem dynamics. Highly impacted and dynamic systems may not respond as quickly or in the ways conservationists might hope, and modelling tools from fisheries science can help to build collaborative understanding of the effects of different scenarios (Tittensor et al. 2018). Studies such as this can be undertaken at a variety of scales to ensure that indicators are not only robust but also contain a full suite of societal preferences, which should make conversations around ecosystem-based management and trade-offs clearer.

6 KEY CONSIDERATIONS FOR AN EFFECTIVE PATHWAY TOWARDS POST-2020 NATURE CONSERVATION

6.1 Introduction

The Convention on Biological Diversity (CBD) was adopted in 1992 and represented the world's first multi-lateral and binding treaty aiming to address the emerging crisis of biodiversity loss. The original aim – to achieve a significant reduction in the state of biodiversity loss by 2010 – was not met (Butchart et al., 2010). Following this, the *Strategic Plan for Biodiversity 2011-2020* and its associated Aichi targets triggered a notable response in protected area policy, climate change agreements, invasive species action and cross-boundary integration. Despite this progress, the signatories' ambitions to halt biodiversity loss will likely suffer the same fate as the 2010 target (Tittensor et al. 2014). While the Aichi Targets may have been intended to be aspirational and different bodies may have different views on their success, the causes of the failure to explicitly meet most of the Aichi Targets are multiple and likely vary across targets. Causes could include, for example, a lack of understanding of the objectives and aspirations of stakeholders (Maxwell et al. 2015), time lags between the implementation of actions and their outcomes (Leadley et al. 2013), the complex and ambiguous nature of the target text (Butchart et al. 2016), and a lack of development of meaningful indicators with which to gauge actual progress made (Hill et al. 2016; Mcowen et al. 2016).

At the 15th Conference of Parties for the CBD in Beijing 2020, governments will negotiate a new biodiversity framework to replace the 2011-2020 Strategic Plan, in alignment with the CBD's 2050 Vision and the 2030 United Nations Sustainable Development Goals (SDGs). A comprehensive and participatory process to develop the post-2020 framework is already underway and will be refined over the next two years, with indications that the post-2020 framework will likely be similar in structure to the existing strategic plan, comprising an overarching goal, strategic goals, targets, indicators and support for national implementation (Convention on Biological Diversity 2019). There are considerable opportunities to translate lessons learnt over the past two decades into meaningful and actionable recommendations for the post-2020 framework. For example, at present, on the ground conservation is not well linked to international goals (Rands et al. 2010) and national policies rarely consider the complex interlinkages and trade-offs governing the relationships between people and biodiversity (Nicholson et al. 2012). Nations must interpret and tailor their approach according to their ecological priorities, cultural underpinnings and socio-economic situations, in a way that is not well supported by the current global biodiversity targets. In particular, as biodiversity transcends borders and is unevenly distributed and impacted, conservation action is not simply additive by individual countries; concerted effort across scales is required in order to reach global-level targets (Arlidge et al. 2018).

A three-day workshop was held in Oxford in July 2018 bringing together academics and conservation practitioners to share lessons learnt and discuss ways forward for international biodiversity commitments. Workshop participants identified three core areas that should be developed as part of the process for developing a post-2020 framework to improve outcomes for biodiversity. They involve 1) Formulating a robust theory of change 2) Integrating modelling and 3) Working collaboratively across scales. These core areas are based on participants' experience in implementing the current Strategic Plan as well as research and practice in other areas of environmental management and conservation. With reference to each consideration, we discuss opportunities for improving the processes around how global targets are set and implemented. Timing is critical as we are now in the position to be able to assess progress retrospectively, prior to the setting of a future agenda for the protection of Earth's biodiversity.

6.2 1: Formulating a robust Theory of Change to link outcomes and actions

Theories of Change (ToC) (Weiss 1997) are crucial conceptual tools to effectively plan and evaluate how desired outcomes are achieved through a series of actions and make explicit the underlying assumptions and risks to the process. They can include a wide range of relationships, influences and pathways as well as feedback loops. A ToC helps to clearly articulate an underlying plan of action which stipulates clear outcomes and the policies necessary to achieve them; without a clear plan underpinning a set of targets, there is a risk of calling for actions that may not effectively lead to the desired consequences. ToC are widely used in international development for planning and evaluating complex challenges, and have seen use in conservation to identify intermediate targets and indicators for monitoring (e.g., Game et al. 2018), and to determine whether conditions and administrative structures are in place to enable the successful implementation of programs or specific interventions (Biggs et al. 2017). Yet none of the headline multilateral biodiversity treaties to date have been explicitly underpinned by defined ToCs or similar. This must be rectified for the post-2020 biodiversity agenda, to ensure that all actors are aware of the rationale behind and links between agreed actions and positive outcomes for biodiversity and society.

The Aichi Targets included a mix of both outcome-oriented and response-based targets, with links between them only identified post-hoc (Marques et al. 2014). This situation meant that there has been confusion about what actions are required towards improving biodiversity. As such implementing policy has been sporadic and ad-hoc, with certain targets gaining more traction than others, meaning the overall aim of the Strategic Plan has not been met (Tittensor et al. 2014; Secretariat of the Convention on Biological Diversity 2016). Formulating the targets within a clear framework, such as by using a ToC, would have helped to increase transparency around the assumptions being made for how each of the targets contributed towards desired outcomes. This is demonstrated in Figure 6-1, where different targets can be linked and shown how they might contribute to the strategic vision of the CBD.

Outcome-oriented biodiversity targets are specified in terms of desired states (e.g. reduce extinction risk of threatened species) rather than action-oriented targets (e.g. improve protected area management). They are necessary to a post-2020 framework as to articulate what change is actually desired as well as act as a reference to consider to what extent that change is being achieved (Collen & Nicholson 2014). Some outcome-oriented targets have been criticised for being overly complex and/or include redundancies and ambiguities that are difficult to operationalize and to ensure consistent interpretation by signatories (Butchart et al. 2016). This may have resulted in a focus on action-based targets which are clearer, can be easier to implement strategies for, and can be measured, such as those regarding protected area coverage (Target 11; 17% of terrestrial and 10% of marine area as PAs) (Jenkins & Joppa 2009; Lewis et al. 2017). However, while terrestrial and marine protected areas have expanded, they do not necessarily cover representative areas of biodiversity, and can still be considered as progress towards the Target even if they are placed outside of ecologically or biologically significant areas (Devillers et al. 2015; Venter et al. 2018). The action-based nature of the target limits its effectiveness and enables countries to claim success while not considering the systems and processes underlying biodiversity loss (Barnes et al. 2018; Jones et al. 2018; Maron et al. 2018).

By structuring targets through a ToC framework , both outcome and action-based targets (interventions) can be included, but structured in a way that makes clear the causal links about how actions contribute to outcomes. Figure 6-1 shows what a ToC might look like for the issue of plastic pollution. Pollution is a key driver of biodiversity loss in both marine and terrestrial systems and is the focus of many existing multilateral environmental agreements, such as the Helsinki Convention and London Convention. Aichi Target 8 reads 'By 2020, pollution, including from excess nutrients, has been brought to levels that are not detrimental to ecosystem function and biodiversity.' In this ToC we show how a target for plastic pollution "By 2050 marine waters are free from plastic pollution" could be worked through at the global scale. This includes working back from the target to workout what outcomes are necessary to achieve the target and finally what interventions might lead to those outcomes. In this case, the ToC approach more clearly guides countries into making commitments to not only cutting plastic pollution but also cleaning up existing and ongoing plastic pollution. By taking such an approach, it becomes much clearer what potential indicators are required to measure various different stages of the system. For example, the outcome target could be measured by plastic levels in the marine environment but to ensure that is reached, monitoring and indicators will be required to measure previous outcomes and also the interventions required.

In the case of the widely lauded Paris Agreement of the UN Framework Convention on Climate Change (UNFCCC), the desired outcome is a maximum average global warming of 1.5 °C. The UNFCCC 1.5°C outcome target has the benefit of clearly defined actions that can help to achieve it; established through years of modelling current and projected emissions from across the world and their effect on future climate, nations can clearly identify and then aggregate the reduction in emissions required to meet this (Rockström et al. 2017). After nearly thirty years of attempting to garner widespread political support for decisive action for biodiversity, momentum is gathering for a similarly clear goal for biodiversity (Watson and Venter 2017). To date, however, biodiversity targets have not been structured with a single overarching outcome target such as degrees warming, but rather aim to capture complexity and scale-dependency through multiple outcome and response targets covering different aspects such as extinction risk, ecosystem services and participation (Mace et al. 2018). While the ToC in Figure 6-1 is somewhat linear, it could be expanded to demonstrate the direction and strength of connections between different actions and outcome targets that would help identify clear opportunities for actions that efficiently contribute to multiple outcome targets. For example, under the Aichi Targets it has been suggested that integrating Target 12 (prevent the loss of threatened species) into spatial conservation planning for Target 11 (protected area coverage) could have led to a fivefold increase in threatened vertebrates adequately covered for only 1.5 times the cost of the cheapest protected area solution (Venter et al. 2014).

The ToC development process can also greatly assist with explicitly planning for target evaluation, something that the Aichi Targets struggled with by relying on evaluation to be implemented post-hoc (Mcowen et al. 2016). The Aichi Targets have been criticised for containing too many elements and those with less elements have found to have seen more progress (Butchart et al. 2016; Green et al. 2019). A ToC approach would help to keep headline targets concise but would allow for additional components to be included. For example, Aichi Target 6 is "By 2020 all fish and invertebrate stocks and aquatic plants are managed and harvested sustainably, legally and applying ecosystem based approaches, so that overfishing is avoided, recovery plans and measures are in place for all depleted species, fisheries have no significant adverse impacts on threatened species and vulnerable ecosystems and the impacts of fisheries on stocks, species and ecosystems are within safe ecological limits." A headline target might be "By 2050, all fish stocks are within sustainable limits" (Figure 6-1). The ToC could then be worked through to include relevant aspects of the target that will lead to achievement of the headline outcome target; in this case underpinned by an ecosystem approach to management.

While there is much attention pinned on the wording of targets to gain action on biodiversity, we contend that a guiding structure will help proper implementation which signatories are currently struggling with (Butchart et al. 2016; Hagerman & Pelai 2016; Sarkki et al. 2016). A ToC would not only be useful from a global perspective (Figure 6-1) but could also be utilised by Parties who would be provided with a clearer pathway of how to translate overall goals to their national context and clearly contribute to the overall vision of the CBD. This would make analysis of commitments more transparent and more straightforward, which could be beneficial to driving negotiations (Parker & Karlsson 2018).

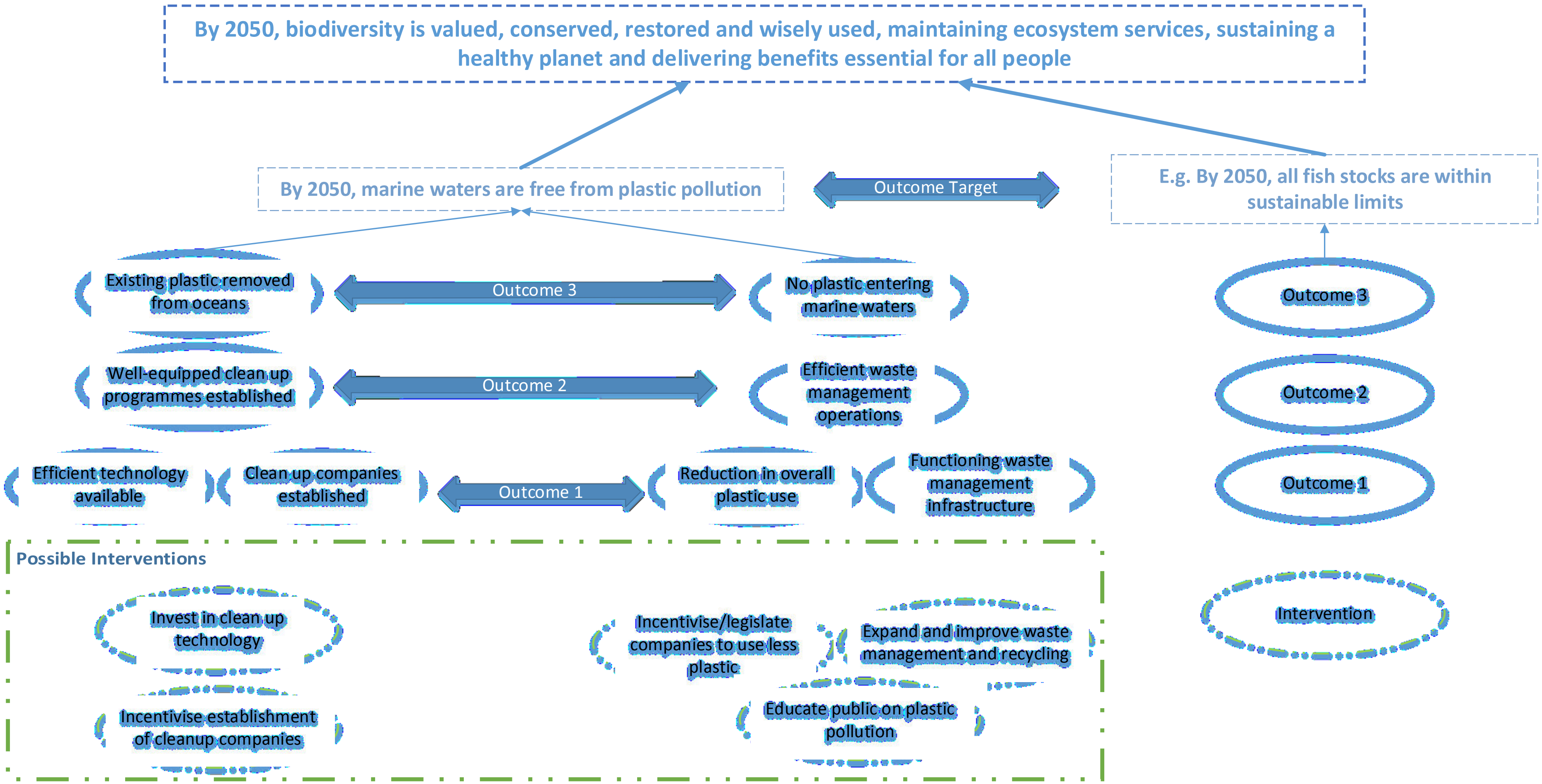

Figure 6-1: Example of how a Theory of Change Model might be created for plastic pollution as part of the CBD Post-2020 Framework. Theories of Change work backwards from deciding a desired outcome and what is necessary for achieving it.

6.3 2: Underpinned by models to integrate complexity and uncertainty

Models are simplified, abstract representations of processes or systems, and can be powerful tools for projecting plausible futures and assisting decision-making at a range of scales (Korzukhin et al. 1996; Starfield 1997). Models range from qualitative conceptual models to quantitative process-based models of dynamical systems and assist with characterization of complex systems and help to constrain and explore uncertainty around future trends. Ecological and socio-ecological modelling have made substantial strides (Nicholson et al. 2019), and underpinning the post-2020 framework with both conceptual and quantitative models can greatly benefit future agreements, as seen by the influence of model-based projections on policy in the climate sphere.

Models provide an explicit qualitative or quantitative description of relationships among various components of systems and can be used to project how such systems might respond to different scenarios. Ultimately the CBD is concerned with ensuring that future decision-making ensures sustainability and enables biodiversity to prosper, thrive and continue to provide the goods and services that human society relies upon. As we can only evaluate plausible futures given large assumptions and specific scenarios, decisions must be robust and adaptable to a variety of different futures under social and ecological change. While models are already used by the CBD to explore different global pathways for achieving its 2050 Vision (Convention on Biological Diversity 2017), they have seen much less uptake and integration in the CBD's 2011-2020 Strategy and in national level policy making. Yet environmental and resource management have used models of varying scales and complexity since the 1970s to inform decision making (Jørgensen 2008). Models have been used to inform conservation and management decisions by predicting future trends and status of biodiversity (Visconti et al. 2016), setting quantitative targets (Desmet & Cowling 2004), developing relevant indicators (Fulton et al. 2005), predicting the likely outcomes of proposed policy or management alternatives (Addison et al. 2013), and to evaluate the effectiveness of those that have been implemented (Law et al. 2017). More recently, highly complex models of multiple processes have been developed to better inform ecosystem-based management approaches (e.g. Fulton et al. 2011).

By projecting forward scenarios under uncertainty, models can aid with science-based target setting, which can help to garner political action. This has been seen for climate change, where models have proved valuable in projecting future climate change and gaining significant progress for action under the UNFCCC aligned to a specific target (van der Sluijs et al. 2010). Model ensembles (e.g. Tittensor et al. 2018) can also provide an exciting and powerful insight into better characterising uncertainty around future trajectories. Advances in model ensembles provides an opportunity for the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) to play a role akin to the IPCC in generating multi-model projections (IPBES 2016).

The use of models alone, regardless of their detail or complexity, cannot guarantee effective conservation actions or success in achieving desired outcomes. However, they can improve the likelihood of success. Where targets have been set outside of a modelling process, models can still be used for policy-screening scenarios and developing counterfactuals to assist decision making (Nicholson et al. 2019). Quantitative models can be particularly useful when combined with qualitative models such as ToC (Point 1). They can assist in validating the assumptions underlying the ToC as well as helping uncover unintended consequences in dynamic feedbacks; global models could be used to project future plastic emissions, which could be used by countries to estimate the necessary capacity increases required in waste management infrastructure (Lebreton et al. 2017). Most importantly models can facilitate quantitative evaluations of potential trade-offs among multiple conservation targets (e.g. between biodiversity and food (Erb et al. 2012)).

While models can help to inform target setting and constrain uncertainty, active and responsive management and evaluation will always be required. Social-ecological systems are highly complex in space and time and not easily translated into quantifiable targets, particularly when stakeholder values are conflicting. Furthermore, even with quantitative targets and relevant indicators, merely focussing on simple monitoring is unlikely to reveal the reasons as to why targets have not been met and how intervention can be improved in the future. The post-2020 framework should ensure that commitments are not only monitored on a target/indicator basis but include provisions for proper evaluation and learning from the very start, which can subsequently feed back into model development and parameterisation. The ever-increasing range of decision science and modelling tools available to scientists and policy makers can help to break down ambiguous, vague or data deficient targets into measurable and or achievable components, both in terms of goals (short- to long-term), and actions (Addison et al. 2013). While Tittensor et al. (2014) provided an early warning that trajectories towards 2020 were not sufficient, it is unclear how this was to be rectified. By including monitoring, evaluation and learning from the start, alongside a ToC approach, global assessment and management of targets can be more dynamic and responsive.

By making assumptions explicit through a ToC (Figure 6-2) and having transparent evaluation plans, opportunity would be created for adaptive management that is responsive and relevant to emerging unforeseen changes. Management strategy evaluation (MSE), for example, is successfully used to manage fisheries, and incorporates multi-stakeholder consultation, modelling, scenario evaluation, and monitoring to allow for structured, adaptive and defensible decision-making (Bunnefeld et al. 2011; Plagányi et al. 2014). Using systems thinking and providing practical guidance to embrace uncertainty and complexity could guide management and support progress towards biodiversity targets. Such an approach has been demonstrated by Stephens et al. (2018), who provide theoretical background and practical tools to consider systems thinking for transformative change in gender equality. . This would increase the flexibility and effectiveness of actions taken and thus the likelihood of their being met, as well as making them more acceptable to stakeholders than rigid targets in the face of uncertainty.

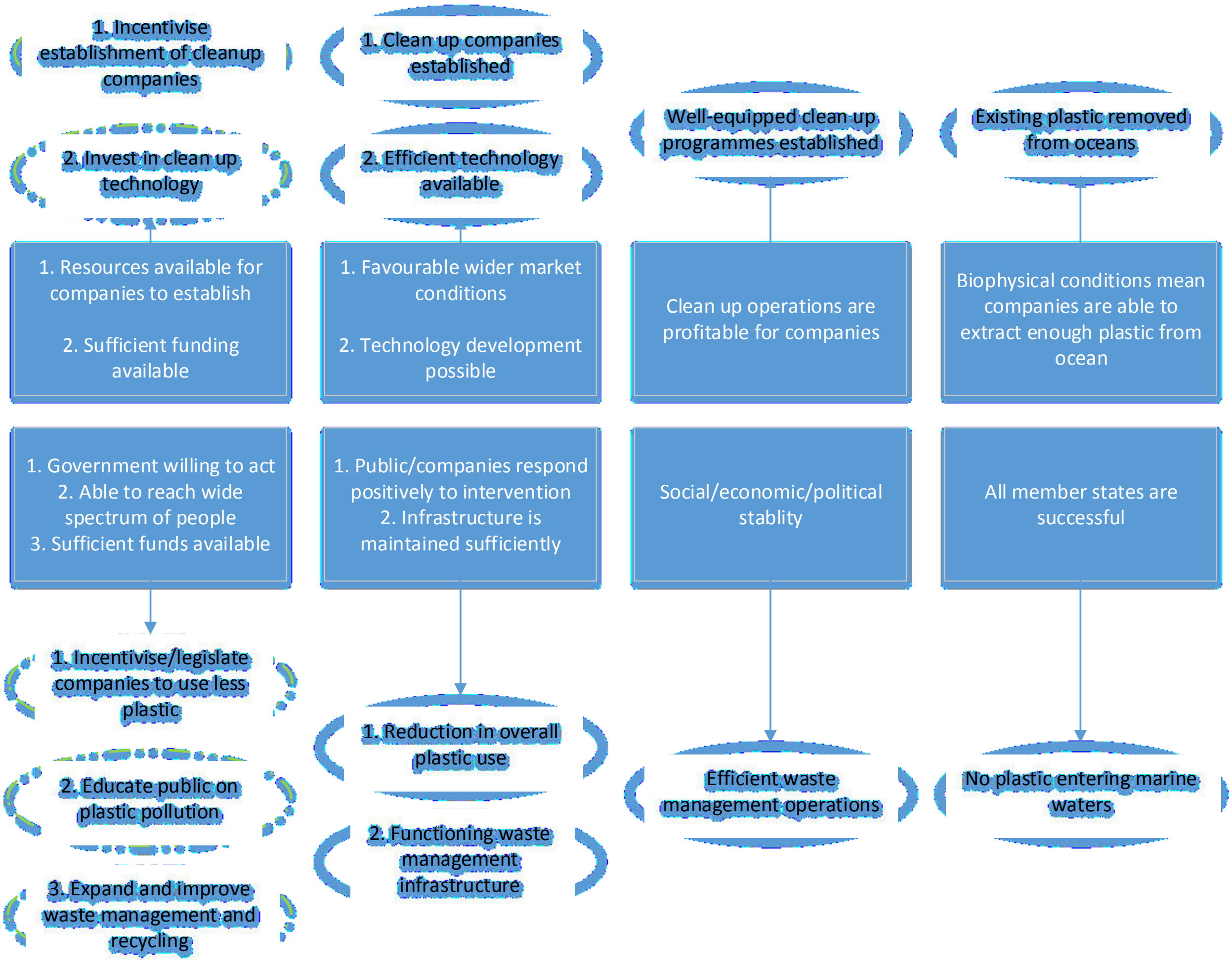

Figure 6-2: Assumptions associated with plastic pollution theory of change. These should be made explicit when undertaking theory of change which helps with planning and transparency.

6.4 3: Transcending scale to inform meaningful devolved and specific local action

Biodiversity, and its benefits to people, are distributed unevenly across the world. For example, 70% of the world's wilderness within national borders is contained within just five countries (Watson et al. 2018). Reaching the Aichi Targets, however, relies on action by all Parties, and some targets (e.g. Target 11) set blanket aspirations regardless of the ecological and economic situation of individual countries (Convention on Biological Diversity 2018b). Countries must agree on an overarching framework but conservation interventions must also be supported by local people to ensure success (Sodhi et al. 2011). Many international treaties have struggled to obtain such buy-in from a grassroots level (Sabatier 1986). Where necessary, the post-2020 framework should transcend scale to help direct differentiated actions at the country level that are most effective towards the vision of the CBD and ensures vulnerable groups are not negatively impacted.

Many biodiversity issues, such as the illegal wildlife trade, are complex and multifaceted, requiring differentiated action from demand reduction at the consumer end, to improved detection and knowledge in transhipment countries, and enforcement and community empowerment at the source (Milner-Gulland 2018); yet such differentiated action is not well supported by current target structure. Other important aspects of biodiversity such as coral reefs or tropical forests are not ubiquitous and require direct local intervention to ensure their persistence, meaning differentiated action is inherently necessary for different countries. Aichi Target 5 requires countries to halve or bring close to zero the rate of habitat loss, but how best to achieve this will depend on the current rate of loss of habitat in different countries, and which habitats are under most pressure. Highly biodiverse countries may favour ensuring critical areas for biodiversity are maintained (Jantke et al. 2018), whereas degraded countries may implement large-scale restoration (Latawiec et al. 2015; van Katwijk et al. 2016). Whilst the CBD does allow flexibility in terms of the opportunity for individual Parties to develop their own approaches (e.g. developing National Biodiversity Strategy Action Plans), relying solely on national actions to achieve global outcomes without clear pathways or prioritisation risks unequal or unjust effort, leading to overall failure to reach global goals (Hagerman & Pelai 2016).

The UNFCCC has been committed to the principle of "common but differentiated responsibilities" since as far back as 1992, although operationalising such a commitment has proved challenging due to disagreement over responsibility (Althor et al. 2016). However, it has provided the opportunity, through Nationally Determined Intended Contributions (NDICs), to identifying gaps between national commitments and the global goal, which is subsequently becoming known as the 'emissions gap' (UN Environment 2018). The CBD pursued an approach relying on national implementation through the National Biodiversity Strategies and Action Plans (NBSAPs) but has been unable to garner the required commitments and countries have struggled with implementation (Hagerman & Pelai 2016; Sarkki et al. 2016). It would be hugely beneficial to the CBD if voluntary actions could be summed up towards overarching global goals; yet this is only likely to be possible, due to the complexity of monitoring and reporting biodiversity, if such action is guided through an overarching framework upfront.

In determining devolved and differentiated activity, such a process would need to ensure equity, which has been challenging for international agreements (Mattoo & Subramanian 2012), but there are distinct advantages for all countries in addressing equity up front (Steffen & Stafford Smith 2013). The UNFCCC Clean Development Mechanism has provided a platform for developed countries to assist with clean development in least developed countries. For the CBD, the Global Environment Facility provides a mechanism for the fair distribution of financial resources to assist countries in implementing the CBD. However, given the complexity of the aims of the CBD, such a process must go beyond simply transferring capital, and drive positive biodiversity enhancement in the developed world as well, based on clear theories of change (Point 1). By blending an overarching ToC with global or regional modelling, more effective mechanisms can be put in place to reach global targets. Figure 6-3 shows a conceptualisation of how the ToC might be operationalised in an equitable manner. While modelling helps assist understanding of where clean up efforts should be targeted, Country 3 is assisted as it does not meet the assumptions listed in Figure 6-2.

While having a shared global perspective is important, top-down approaches commonly ignore or inadequately incorporate local perspectives, particularly those of vulnerable groups. This can result in both failed biodiversity outcomes and costs to poor people (Franks et al. 2014). Global-scale targets often ignore, or even consciously limit, local community participation despite this being a critical component for on-ground conservation success (Phelps et al. 2010). Advancing the current understanding of location, cultural diversity, and scale-dependent effects of the present biodiversity crisis is a major challenge to carrying out effective management actions (Garnett et al. 2018). The CBD has gone substantially further than many other international treaties by providing frameworks to engage directly with Indigenous Peoples and Local Communities at the level of the Secretariat. However, the CBD is primarily implemented through the NBSAPs at the national level, where there has been little meaningful engagement and inclusion of these groups (Cooney et al. 2018). In an increasingly modernising world, the power of the public should not be understated, as seen in the recent action on plastic pollution (Kontrick 2018) and increasing interest in the environmental impact of peoples' diets (Poore & Nemecek 2018). Exploring novel ways of understanding impacts and actions and helping people to understand how they can best minimise their impact on biodiversity (informed by a global outlook) would help connect people not only to the vision of the CBD but potentially improve outcomes for biodiversity through consumer action and ultimately political pressure (Dorward et al. 2017).

Finding a balance between global efforts to advance effective action in the right locations and local efforts that allow communities and nations to use the resources needed to develop economically is a key challenge. Improved guidance about how actions lead to outcomes (Point 1) and addressing issues of equity at the country and local level can allow for better alignment of priorities across scales. Informed by scenario modelling where possible (Point 2), these factors would allow countries to contribute targeted, effective actions towards global biodiversity outcomes and ensure biodiversity is not just the responsibility of those countries that contain intact or unique biodiversity.

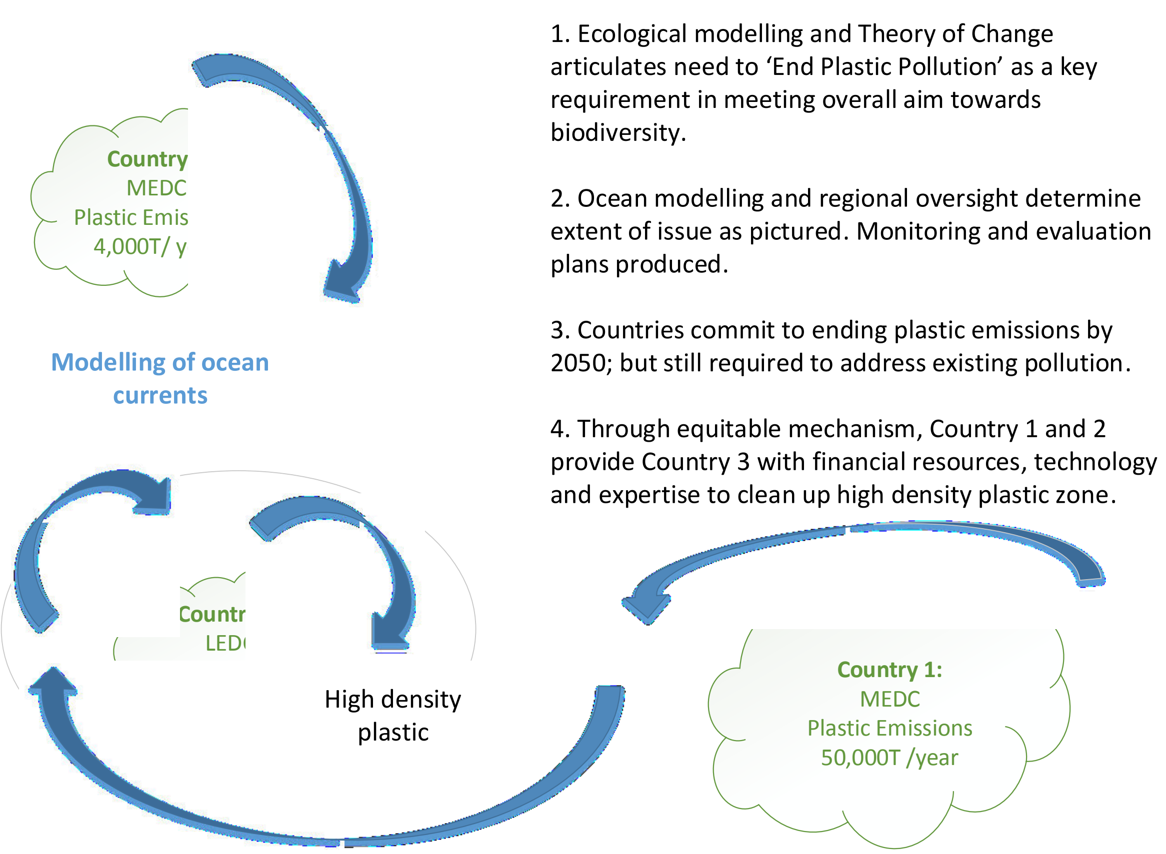

Figure 6-3: Conceptualisation of key recommendations in stylised example on plastic pollution. This describes three countries and how plastic pollution could be best tackled from utilising theory of change, using modelling studies and transcending scale from regional to local.

6.5 Conclusion

Global biodiversity is declining. If we are serious about protecting and restoring biodiversity, then actions will need to be wide-reaching in scope and geography – merely formulating new targets as updates of the existing targets is unlikely to create meaningful change. It is essential that the post-2020 framework focuses not only on what needs to be done but also how it should be done, using measurable steps which make sense at the scales at which biodiversity change happens. A Theory of Change model would provide a useful overarching framework at the global scale to link outcome and action-based targets in to a transparent format that would help Parties implement the overall vision of the CBD. Such a framework could integrate recent advances in modelling and decision science to ensure that system links are explored, interventions have a sound basis and evaluation is planned up front and in a way that goes beyond simple target/indicator relationships. Importantly, a framework would also need to support differentiated action at a range of scales to ensure mechanisms are put in place that can assess how national contributions scale up to global outcomes, whilst protecting the most vulnerable in society.

7 DISCUSSION

7.1 Introduction

The need for and use of indicators at large scales is unlikely to slow. Indeed, as international agreements such as the SDGs progress and start to be evaluated, indicators will come under increasing scrutiny (Hák et al. 2016b). While the literature is extensive on what constitutes good indicator design (Failing & Gregory 2003; Niemeijer & de Groot 2008a; Jones et al. 2011), there remains little formal evaluation of the context and usefulness of existing indicators (Bockstaller & Girardin 2003; Moriarty et al. 2018). As such, when indicators receive criticism or fail to make an impact, new indicators are regularly proposed (e.g. Butchart et al. 2005; Scholes & Biggs 2005; Collen et al. 2009; Certain et al. 2011; Halpern et al. 2012; Hsu et al. 2014), often in the hope of finding the overarching 'Gross Domestic Product' (GDP) for biodiversity (Balmford et al. 2005). This thesis aimed to examine existing large-scale indicators and to see how progress can be made around their design and use going forward, rather than searching for or promoting the "GDP" indicator for biodiversity.

While this thesis addresses various elements of indicator design and validation, three core overarching themes emerge. The first is an examination of the role of composite indicators; what are their strengths and weaknesses, how can they be validated and what is their place in the context of large-scale national and international policy processes. The second theme is how we can progress model-based testing and development of indicators. Whilst testing indicators was seen as a priority for the Convention on Biological Diversity (CBD) it has largely not been enacted (Convention on Biological Diversity 2016). The third theme is how indicators can be used across different scales. This thesis has contributed to better understanding in these three areas, and I discuss each of these in turn here, while also highlighting areas for future research.

Subsequently, I briefly discuss another fundamental question, which remains open and seemingly critical. That is, how do we contextualise, develop and use indicators within large-scale national and international policy decisions. Even with enhanced testing and validation, the structure and context that indicators are selected and used in must be fit for purpose.

7.2 Overview of thesis

The aim of this thesis was to investigate the challenges related to the use of large-scale environmental indicators in marine conservation and explore how these can be addressed. In order to achieve this aim, I have explored both a structured approach through composite indicators and a systems based approach drawing on modelling techniques developed in the Nordic and Barents Seas. In Chapter 2 I review the literature to explore the full suite of uncertainties associated with composite indicators throughout their whole life cycle, before summarising key methods to address them. Chapter 3 takes the structured approach of the Ocean Health Index (OHI) to the Arctic Ocean region, highlighting its value in pulling together disparate data and making a high-level assessment of ocean health at the regional scale. Drawing on Chapter 2, in Chapter 4 I take a systems based approach to indicator testing. I use a social-ecological model to look at responses of two well-known biodiversity indicators and find that construction methods can impact indicator performance, by reducing sensitivity through aggregation and sub-indicator selection. Chapter 5 then projects fishing scenarios into the future with climate change to see how a both fisheries and marine conservation indicators respond. They differ in that while fisheries indicators point to improved biodiversity performance, conservation indicators report on biodiversity decline. I show how model-based projection can help to interpret indicators and aid decision making. Finally in Chapter 6, I look at the context in which indicators are used and make recommendations for how the CBD approach could benefit from a more transparent framework which integrates models, guides action and works across scales.

7.3 Core themes across thesis

7.3.1 The role of composite indicators

Composite indicators have continued to gain popularity in a range of sectors for having the ability to display and compare multiple different but related attributes (Paruolo et al. 2013). Despite their popularity, they have been criticised for an under-consideration of uncertainty and their dynamics not being full explored (Böhringer & Jochem 2007; Jørgensen et al. 2013; Giampietro & Saltelli 2014). While there has been significant attention placed on some methodological aspects, particularly weighting and aggregation, this thesis has shown that careful attention and thought is required throughout all stages of composite indicator development (Chapter 2). The attention placed on mathematical considerations in composite indicators may be due to the many composite indicators in different fields, such as technology, tourism or transport, which purposely measure contrasting components that have little interaction (Grupp & Mogee 2004; Famurewa et al. 2014; Mendola & Volo 2017). However, composite indicators used within environmental science such as the OHI (Halpern et al. 2012) or the Environmental Performance Index (Hsu et al. 2013), tend to measure aspects of systems that are highly interrelated. Chapter 2 illustrates that the process behind developing the theoretical framework and the dynamics underpinning indicator development are less considered, or not easily communicated or understood, which can have a detrimental effect on validation. As such I recommended systems modelling as a key method to advance consideration of systems dynamics within composite indicator design and use.

In Chapter 4 I take the ideas of systems modelling in Chapter 2 into a demonstration of how composite indicators can be integrated within a systems modelling approach in the specific setting of the Barents sea. By exploring systems dynamics within the model, I show how the selection of species and the level of weighting and aggregation actively constrain the sensitivity of the Norway Nature Index (NNI). This is somewhat concerning as the management intervention in this example was relatively extreme; it applied heavy fishing pressure to an unfished system, where we would expect to see large indicator changes. In conditions more akin to real life, in Chapter 5, we see that the NNI shows little variation between scenarios, which would make its interpretation more difficult without a modelled understanding of the impact of management on the different interacting components of the system. Future research here could build on this work to re-think what species are included within the NNI and how they are weighted. By modelling the ecosystem interactions, it is clear that some species are not well suited to indicator inclusion; for example some species (e.g. zooplankton) have extremely high inter-annual variability, which clouded the signal of the pelagic NNI and subsequently the overall NNI. The NNI is an index which has purposefully been created to measure Norway's biodiversity performance over time. As such it is important that the index has the ability to respond to system change if it is to be useful to decision making.

While systems thinking within composite indicators remains underdeveloped, one of their largest assets is in providing a structured framework that can be clearly communicated. Chapter 3 showed the value of taking a structured composite indicator approach to a large data-poor area, the Arctic. Pan-Arctic studies are limited due to the complexity of compiling data across multiple regions which have different monitoring regimes and are often data-limited due to harsh climatic conditions (Hamilton & Lammers 2011). I provided the first attempt at compiling and analysing data on a range of aspects of ocean social-ecological systems in the region. The OHI Framework facilitated this compilation. Although the Arctic is performing reasonably well in many goals, this is not evenly distributed across countries or sectors. Biodiversity-focussed goals performed well showing how improved ecosystem management through recovering fisheries and sustainable marine mammal exploitation were having a positive effect. Conversely, other goals showed that there was significant room for improvement; particularly in sustainable tourism, mariculture, fisheries, and protected places. Unified assessments such as this one can support national comparisons, data quality assessments and discussions on the targeting of limited monitoring capabilities at the most pressing and urgent transboundary management challenges, which is a priority for achieving successful Arctic stewardship. For example, I found that data on marine mammals was particularly limited, despite playing a key role for food provision and resources such as ivory and pelts, as well as many being globally iconic and regionally and locally culturally important (Laidre et al. 2015). Many marine mammals have large ranges and are shared resources across the Arctic. Greater understanding of their status across their range, how this might be altered under climate change and where and how they are used by people would assist with their regional management. The Arctic Council already provides a structure for regional decision making by convening the eight Arctic nations and six indigenous peoples' councils. However, it takes a specialist working group and task force approach to individual issues such as biodiversity, circumpolar monitoring and pollution, which remain relatively siloed (Koivurova 2010). By taking a structured overarching approach, such as through an OHI assessment, the linkages and complexities across social-ecological systems in the region may be better explored at higher policy levels.

While significant effort has gone into developing multi-dimensional indices such as the OHI for potential use with global-scale policy, their future use in this domain appears unlikely. None of the 230 indicators used to measure the SDGs are multi-dimensional; even long-standing indices such as the Human Development Index have not been included (Rickels et al. 2016). Indeed only five of the indicators included are indices; although perhaps interestingly three of these are included within SDG 14 and 15 (Red List Index, Mountain Green Cover Index and an index of pollution), showing that conservation has a tendency to lean towards aggregations of indicators (United Nations Statistical Commission 2017). However, as I show in Chapter 4, there is a disconnect between global and national level indicators. The development of the OHI has already proved useful as a way of structuring thinking around ocean health in many countries across the world and is likely to continue to do so (Halpern et al. 2013b; Elfes et al. 2014; Selig et al. 2015; Daigle et al. 2017). The requirements on countries to work towards

SDG 14, let alone all of the SDGs, require complex decision making. Composite indicators such as the OHI can provide a solid starting point for considering multiple and competing aspects around sustainable ocean development. Given the interconnectedness of marine systems, the OHI may be most usefully applied in a regional approach as demonstrated here (Chapter 3) and in the Antarctic (Longo et al. 2017). Many aspects such as fisheries, which require assessment at the stock scale, or pollution, often exceed national EEZ boundaries and therefore require regional management. The Baltic Sea has begun its own OHI as a way to bring together stakeholders across the region to manage an area that has historically suffered the effects of poor transboundary management (Elmgren et al. 2015). The importance of taking regional and transboundary approaches to ocean issues is clear, yet is not well supported by international processes such as the SDGs or CBD; it is here where the Ocean Health Index may find its most appropriate role.

7.3.2 Systems Modelling and indicators

In order for indicators to be useful within national and international decision-making for biodiversity it is important to understand not only how to 'bend the curve' but if the indicators themselves will actually respond as predicted (Mace et al. 2018). Testing and developing indicators was a priority for measuring progress towards the CBD 2010 target to ensure indicators could effectively report progress and communicate trends in biodiversity; however 40% of indicators were still listed as requiring further testing (Convention on Biological Diversity 2006). Despite key studies highlighting the possibility and importance of using systems models to test biodiversity indicators (Jones et al. 2011; Nicholson et al. 2012; Costelloe et al. 2015), this was likewise not properly enacted for the current CBD Strategic Plan to 2020 (Convention on Biological Diversity 2016). Testing may have been downgraded as a priority in order to merely find indicators that fit, as three of the 20 Aichi Targets were still without indicators entirely in 2016 (Mcowen et al. 2016). In Chapter 2 I explore the basis of systems modelling for composite indicator testing and development. I explain that where possible, taking a systems-based approach first can be extremely beneficial in guiding indicator selection but also for weighting of indicators and understanding what data gaps are present. In Chapter 4 and 5 I then highlight two examples of the benefits of taking a systems modelling approach towards indicator validation.

By using a systems approach in Chapter 4, I was able to look at how the Living Planet Index (LPI) and NNI respond to fairly extreme system change. The NNI was constructed in my model similarly to its real-life counterpart as many of the species selected for the NNI are present in the model. The LPI, however, relies on all available data and therefore contains a wide range of different species that are not included within the model. In our case therefore, the test of the LPI is of the methods used for its construction, and not of the data underlying the real-life LPI. Being unable to exactly reproduce the LPI, however, should not be an argument against development of an approach to testing it. If we look at the lessons from Chapter 2, we see that there is a great advantage in starting indicator design and selection from a systems based approach. If we are truly interested in indicator validation, which has been seen as extremely important by the CBD and wider literature (Collen & Nicholson 2014; Moriarty et al. 2018), then better linking of indicators and models, as I have done here, should be considered a key step forward. Fisheries science has widely embraced these techniques, resulting in the discrediting of one of the most widely used indicators, mean trophic level (MTL) from catches. This indicator intends to detect shifts from high-trophic-level predators to low-trophic-level invertebrates and plankton-feeders, in what has been described as “fishing down food webs” (Pauly & Watson 2005). However, through modelling approaches it was found that fisheries collapses can occur even when MTL is stable or increasing (Branch et al. 2010). It has since been removed from the CBD indicators list (Convention on Biological Diversity 2016).

The existence of marine models across the world offers a starting point for immediate furthering of testing species-based indicators. Testing the Red List Index (RLI) could be an important next step, considering its prominence in both the SDGs and Aichi Targets, and its potential issues with taxonomic bias and responsiveness (Costelloe et al. 2015). The RLI requires species-level data and therefore is most appropriately tested at the global scale. While there is not currently a mechanistic global ocean model, combining different regional models may offer a way forward (Olsen et al. 2018). Following the methods laid out in Chapter 4 and 5, the LPI could be further tested in many modelling frameworks. This would also have the advantage of better integrating databases used for modelling within the global LPI, which would boost its overall coverage; many of the species present within the NoBa model are not present within the global LPI database. If model data and the global datasets were better integrated, further research could disaggregate the LPI for species included within models and compare these outputs to those obtained when including all species; this would help understand how model-based indicators compare to those in the real world which are based on a much wider array of data (Fulton et al. 2005; Nicholson et al. 2012).

Indicators help communicate complex aspects of biodiversity change to wider audiences, such as policy makers. The use of models and scenarios largely remains a scientific endeavour, with weak uptake by policy makers (Nicholson et al. 2019). Part of this may be the complexity of the models themselves and of the communication of their results. If there were greater confidence and consensus on indicators through their integration with models, then scenario projections could be communicated in a way that would resonate with policy makers and wider stakeholders. In Chapter 6 I explain how models and scenarios can underpin global biodiversity targets to help understand complexity and constrain uncertainty. While modelling approaches are well developed and integrated into climate policy, biodiversity models have received much less attention. Climate modelling has allowed the clear articulation of a maximum average global warming of 1.5 °C, which is both science-based and clearly communicated. If biodiversity is to find a similarly uniting global target, then a single or small number of indicators will be required to measure progress and analyse pathways. This will require advanced testing and integration with models.

In this thesis I focus on marine conservation, where models have been developed to inform fisheries management (Fulton et al. 2011a) and are constantly evolving and improving. Multi-model comparisons of Atlantis models and other modelling frameworks has been an exciting development in exploring and potential future scenarios of change across different systems and seeing how they compare (Olsen et al. 2016, 2018), while cross-model comparisons are quickly emerging as ways of comparing how different models perform (Tittensor et al. 2018). Such projects will significantly increase the quality and power of modelling for marine systems going forward. Terrestrial systems are modelled less widely in mechanistic sense, with terrestrial modellers favouring statistical approaches (e.g. Newbold et al. 2015); but these cannot generally capture or predict non-linear and dynamic responses to perturbations (Evans et al. 2013). A reason for this is potentially because terrestrial dynamics are less clear than in marine systems and that general rules cannot be abstracted across ecosystems (Purves et al. 2013). However, there appears to be a growth in terrestrial modelling, through initiatives such as the Madingley model, that can build upon these advances in marine science (Bartlett et al. 2016).

Dynamic models are far from infallible, however, and their weaknesses as well as their strengths must be considered. End-to-end models, such as Atlantis which was used in this thesis, are highly complex and data intensive. Atlantis models have been developed in 35 regions around the world, from small estuaries to large ocean regions. A good deal of experience, training and guidance is necessary to develop Atlantis models and a single run from the NoBa Atlantis model takes around 13 hours to process, so time and resources must be carefully considered (Hansen et al. 2019). Many assumptions are made during model development about how to represent processes and deal with data gaps; in itself this is not inappropriate for theoretical exercises such as those contained within this thesis, but caution must be taken if utilising such models for real-world decision making (Fulton 2010). Likewise, few models are formally validated to understand how well a model reproduces a true system state. Undertaking such a process is known as skill testing and is typically started during model development by hindcast skill assessment to see how well a model is able to replicate historical data. However, a model that has high hindcast skill will not necessarily have high forecast skill when confronted with data outside of that used for parameter estimation (Francis et al. 2011). Forecasting, such as applied in Chapter 5, is important for informing decision-making. Determining model forecast skill involves using the model to forecast future conditions, and then, in the future, evaluate how the model predictions compare with the subsequently available observed data(Olsen et al. 2016). We are now in a situation to forecast skill test some models and as time goes on, this will become more widely applicable.

While such limitations may act as a barrier to increased uptake of such complex modelling approaches, other less complex models may well have great value for indicator development and testing (as discussed in Chapter 2). Given the current lack of testing, even simple examples, such as in Costelloe et al. (2015), who do not consider systems dynamics, allow for some exploration of indicator performance. Models of intermediate complexity (MICE) and minimally realistic models offer interesting potential for more tactical use of models to answer specific questions and have in some situations been seen to work better than end-to-end models (Plagányi et al. 2014). MICE estimate parameters through fitting to data, use statistical diagnostic tools to evaluate model performance and account for a broad range of uncertainties. They are much smaller, faster and more easily interpreted than counterparts such as Atlantis, and best deployed in specific contexts. They limit complexity by restricting their focus to those components of the ecosystem needed to address the main effects of the management question under consideration. MICE could potentially work well with more specific indicators or at smaller scales, such as those that measure trends in threatened or iconic species at national levels to feed into the successor of Aichi Target 12. MICE have already been used to quantify the impacts of whaling for example (Tulloch et al. 2018). Likewise in cases where quantitative models are not available, conceptual models can provide a basis for understanding linkages within systems and how indicators can best be developed and used (Rowland et al. 2018). Such an approach would have been useful to combine with the OHI assessment of the Arctic in Chapter 3. While the structured framework of the OHI is useful for drawing data together and making preliminary assessments, any movement towards specific changes in management would need to consider the system as whole and how various factors trade-off. This could potentially lead to alteration and iterative improvement of the OHI developed for Chapter 3.

7.3.3 Using indicators across scales

How actions, information and outcomes differ and change across scales has cut across almost all chapters in this thesis. As I highlight in Chapter 6, in order to achieve global biodiversity outcomes, the issue of scale requires greater consideration and clarification. While global leadership is undoubtedly important and necessary, actual action and change comes from nation states, who therefore need to work together to scale up to global outcomes. There is a distinct lack of clarity in the current CBD system on how to scale back and forth between international commitments and national-level actions, and analyses of progress to date have shown nations are not delivering enough conservation action to meet global goals (Tittensor et al. 2014; Secretariat of the Convention on Biological Diversity 2016). Part of the problem with this 'biodiversity gap' is that it is measured post-hoc, rather than forward planned (Chapter 5 and 6).

Measuring the 'biodiversity gap' post-hoc is not straightforward as indicators and data are not connected between the national and global scales. While the CBD does not prescribe what indicators should be used for national reporting, it does require its global-level indicators to be able to be used at the national scale (Convention on Biological Diversity 2016). In reality, as we highlight in Chapter 4, that means data that could be used for indicators is being lost as bottom-up data isn't feeding into global indicators. This is also seen in Chapter 3, where the data and models produced for the Arctic OHI are not comparable to the global or other regional or national OHIs. While initiatives such as Essential Biodiversity Variables attempt to level the playing field in terms of providing detailed data at a range of scales, they do not necessarily link in well with existing indicator efforts, such as the LPI and RLI, which have already undergone significant development (Pereira et al. 2013). A key finding of Chapter 5 is how indicators attempting to measure similar aspects of system change and using the same data can perform differently and tell different stories. This is likewise true across scales, where global-scale indicators can hide regional differences (Hill et al. 2016). By testing indicators across a range of scales, an effective example of which I show in Chapter 4, confidence in indicators can be built and the data required to improve them can be sought. Likewise, a demonstrable and understandable link can be made between indicators that are used at different scales to inform policy decisions. For example, in Chapter 5 I show that changes in fisheries management have a minimal impact on future biodiversity as indicated by fisheries ecosystem indicators, but conservation indicators point towards overall downward trends in all scenarios because of climate change, which would require additional conservation action outside fisheries management. These comparisons are difficult and often masked at higher scales and therefore must be exposed at smaller spatial scales, requiring systems modelling and indicator testing to be advanced at more local scales such as in Chapters 4 and 5.

The need for scaling up national actions to global outcomes (Chapter 6), will probably require consolidated effort at regional scales. This is particularly apparent for ecosystems which are unique or irreplaceable, such as the Arctic. The Arctic is a particularly interesting case as it contains globally iconic species and fragile habitats and is highly threatened by global forces through climate change (Harris et al. 2017). In Chapter 3, I find Arctic monitoring and management to be disjointed and poorly integrated, which threatens the overall stability of the region. While a key highlight of this chapter was demonstrating a framework that can start to think about both social and ecological management and trade-offs for this region, management actions would need to be fed by an understanding of what would be most effective in reaching global goals. In this region, this understanding would probably need to be obtained through conceptual modelling (Chapter 2), but quantitative models as shown in Chapter 4 and 5, could feed in to the process to help parameterise parts of a conceptual model. Climate and sea ice modelling for the region is particularly advanced and a key driver of change; these quantitative models could help advance and parameterise future conceptual models around ecosystem and social change by predicting the extent of the main driver of change in the region (Crépin et al. 2017). Through bodies such as the Arctic Circle, it could then be clearly communicated what external or global action is required to meet regional objectives, which could be better integrated into global goals (Chapter 6).

7.4 The future for indicators

This thesis largely focusses around the methodological aspects of indicator design and use from a scientific perspective; highlighting issues with indicators in their current format and demonstrating potential methods to constrain uncertainties and integrate complexity. There is a huge assumption within the conservation community that more specificity and greater quantification of targets will lead to more and better action and thus better outcomes for biodiversity (Butchart et al. 2016; Green et al. 2019). This may help in progressing the issues related to sufficiency and suitability of indicators to measure targets, assuming that quantification of the target can actually be expressed with an indicator (Mcowen et al. 2016), but at the same time risks becoming subject to Goodhart's law, where a measure ceases to become a good measure when it becomes a target; for example using the RLI as an indicator may drive action solely towards the most threatened species, while overall biodiversity is neglected (Newton 2011). It also fails to recognise the many auxiliary functions that diverse global biodiversity targets have, such as raising awareness, building partnerships, promoting investment and developing tools and knowledge (Doherty et al. 2018). To influence target setting, conservation scientists need to be able to propose targets that they are confident will lead to better outcomes for biodiversity (Barnes et al. 2018), while being sensitive to trade-offs with other aspects of social-ecological systems and local requirements, such as food provision, livelihoods and poverty (Singh et al. 2018).

In Chapter 6 I build on the lessons learnt in the earlier parts of the thesis, to think about processes that can help to drive better biodiversity outcomes, rather than the targets themselves. The challenges faced around indicators are not merely scientific; indicators are closely linked to targets that are politically negotiated, with little consideration of how they might be interlinked or measured (Maxwell et al. 2015). Conservationists therefore should be more focussed on informing the processes behind which targets are formulated and actioned, to ensure that countries can be informed about trade-offs or win-wins for biodiversity. This will require more effort by conservationists to use models at a range of scales (as in Chapter 4 and 5), to test and compare indicators and explore different futures (Wood et al. 2018). While models and projections can be used with indicators on a purely scientific basis, such as in Visconti et al. (2016), if we are to make progress within international forums, the science involved has to be seen not only as credible, but also as salient and legitimate, working within the context of these international agreements and with a range of stakeholders (McNie 2007).

Chapters 4 and 5 in this thesis are purposefully constructed around using existing models and indicators to provide salience. Proposing to rewrite decades of progress by requiring new models and indicators could be extremely damaging and counterproductive. Legitimacy was added to the process through using existing models and indicators in Norway by working with both the Norwegian Institute for Nature Research (NINA) who developed the NNI and the Institute for Marine Research (IMR) who developed the NoBa model. Chapters 4 and 5 have provided a building block for Norway to consider how it uses information to make decisions and I am currently talking to the Norwegian Institute for Nature Research (NINA) about assisting on a project to further test and develop the NNI. This work has helped to build a bridge between the two organisations with regard to the NNI. Expanding similar work to other countries could improve decision-making not only at the national scale but provide countries with the knowledge of what is required at the global scale.

I hope that the work in this thesis can be used positively towards supporting the CBD, and other large-scale national or international policies and agreements, by advancing the conversations around indicator design and use, particularly across scales. The findings of this thesis suggest that conservationists and modellers need to work more closely together and that is possible to do so. By integrating conservation values into modelling processes, better decisions can be made that factor in competing and conflicting objectives.

7.5 Recommendations and lessons learnt for indicators

This thesis included a diverse range of aims exploring the design and use of different indicators for use in marine conservation. These included the exploration of uncertainty within composite indicators, the use of modelling frameworks for informing indicator validation and practical usage, as well as assessing how indicator usage can be improved in future international agreements. From undertaking this research, a number of key recommendations and lessons learnt have become apparent for both future research and practice in relation to indicators:

- While indicators are often the subject of intense scrutiny and debate, it is important to remember that they often serve two quite distinct purposes, which can often be in conflict; communication and management. The communication side is often linked to simplifying an output so that it is widely understood, particularly by non-specialists. Indicators are also used for management and decision making, which often requires more detail or information than an indicator can present. Having greater awareness of this and the limitations of each option, may assist in helping to consider what the indicator is required for and how will it be used and not just what the indicator should be.
- To be truly 'useful' indicators must be linked to some form of management and ideally should indicate towards some kind of management objective or desired state. In an

ideal world, trigger points would be set so that management interventions would change when indicators hit certain points of thresholds. In order to do this well, indicators should be linked to conceptual systems models and reflect the parameters and components within this. Models can then act as the basis for both selecting indicators and testing their performance. Theory of Change models can provide a powerful and transparent basis for working through management problems and selecting meaningful outcome indicators.

- This thesis dealt in detail with composite indicators and a number of recommendations specifically for their design and use can be made:
    - Composite indicators often include a mixture of different types of indicators, for example the OHI attempts to indicate both present state and likely future status (based on an extrapolation of trend). State and rate are different concepts and care should be taken if merging these within a single index. Ideally this should be avoided, but if necessary it should be clear how they relate and how they are interpreted. For example, increase in the rate of extinction risk could be the same for two species, but would be interpreted very differently if one species was close to extinction and one was not.
    - Weighting and aggregation of composite indicators should be carefully considered to avoid arbitrariness, eliminate redundant or highly correlated aspects and to ensure the index is efficient, relevant and easily interpreted. Geometric aggregation appears to be a straight forward and key way of helping in this regard.
    - Indicators are often scaled around min/max values, which if exceeding the max value, the indicator achieves perfect score. Careful consideration is required for this approach to ensure the max value is reliable and that the indicator is responsive to changes.

7.6 Concluding remarks

As 2020 approaches, the CBD renegotiations are coming closer and analysis turns towards understanding the factors underpinning progress towards the SDGs. In the next year, therefore, the use of indicators will come under increased scrutiny and further debate. This thesis contributes to the growing body of literature in this domain by exploring different types of indicators, their uncertainties, their usefulness and techniques for their validation and testing. Finally, it reflects on the role of indicators within the bigger picture and how structured processes can improve how indicators are designed and used. While there is much work to be done, further work should be done practically, alongside government and international agencies in order to create a better system for designing, monitoring and reporting upon environmental management. Scientists cannot change the status quo by merely discussing problems, but must now provide solutions in a positive and practical manner, with a wide range of stakeholders.

8 REFERENCES

Addison, P.F.., de Bie, K. & Rumpff, L. (2015). Setting conservation management thresholds using a novel participatory modeling approach. *Conserv. Biol.*, 29, 1411–1422.

Addison, P.F.E., Rumpff, L., Bau, S.S., Carey, J.M., Chee, Y.E., Jarrad, F.C., McBride, M.F. & Burgman, M.A. (2013). Practical solutions for making models indispensable in conservation decision-making. *Divers. Distrib.*, 19, 490–502.

Agardy, T., di Sciara, G.N. & Christie, P. (2011). Mind the gap: Addressing the shortcomings of marine protected areas through large scale marine spatial planning. *Mar. Policy*, 35, 226–232.

Agyemang, I., Mcdonald, A. & Carver, S. (2007). Application of the DPSIR framework to environmental degradation assessment in northern Ghana. *Nat. Resour. Forum*, 31, 212–225.

Althor, G., Watson, J.E.M. & Fuller, R.A. (2016). Global mismatch between greenhouse gas emissions and the burden of climate change. *Sci. Rep.*, 6, 1–6.

AMAP. (2017). *Snow, Water, Ice and Permafrost in the Arctic (SWIPA) 2017*. Oslo, Norway.

Andersen, K.H., Jacobsen, N.S., Jansen, T. & Beyer, J.E. (2017). When in life does density dependence occur in fish populations? *Fish Fish.*, 18, 656–667.

Anderson, S.C., Cooper, A.B., Jensen, O.P., Minto, C., Thorson, J.T., Walsh, J.C., Afflerbach, J., Dickey-Collas, M., Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Mosqueira, I., Rosenberg, A.A. & Selig, E.R. (2017). Improving estimates of population status and trend with superensemble models. *Fish Fish.*, 732–741.

Arctic Council. (2011). Peoples of the Arctic [WWW Document]. URL http://www.arctic-council.org/index.php/en/environment-and-people/arctic-peoples/122-peoples-of-the-arctic

Arlidge, W.N.S., Bull, J.W., Addison, P.F.E., Burgass, M.J., Gianuca, D., Gorham, T.M., Jacob, C.D.S., Shumway, N., Sinclair, S.P., Watson, J.E.M., Wilcox, C. & Milner-Gulland, E.J. (2018). A Global Mitigation Hierarchy for Nature Conservation. *Bioscience*, 68.

Arrigo, K.R., Perovich, D.K., Pickart, R.S., Brown, Z.W., Dijken, G.L. Van, Lowry, K.E., Mills, M.M., Palmer, M.A., Balch, W.M., Bahr, F., Bates, N.R., Benitez-nelson, C., Bowler, B., Brownlee, E., Ehn, J.K., Frey, K.E., Garley, R., Laney, S.R., Lubelczyk, L., Mathis, J., Matsuoka, A., Mitchell, B.G., Moore, G.W.K., Ortega-retuerta, E., Pal, S., Polashenski, C.M., Reynolds, R.A., Schieber, B., Sosik, H.M., Stephens, M. & Swift, J.H. (2012). Under Arctic Sea Ice, 336, 2012.

Aslaksen, I., Nybø, S., Framstad, E., Garnåsjordet, P.A. & Skarpaas, O. (2015). Biodiversity and ecosystem services: The Nature Index for Norway. *Ecosyst. Serv.*, 12, 108–116.

Athanasoglou, S., Weziak-Bialowolska, D. & Saisana, M. (2014). *Environmental Performance Index 2014 JRC Analysis and Recommendations - JRC Science and Policy Reports*.

Audzijonyte, A., Pethybridge, H., Porobic, J., Gorton, R., Kaplan, I. & Fulton, E.A. (2019). Atlantis : a spatially explicit end-to-end marine ecosystem model with dynamically integrated physics, ecology and socio-economics modules. *Methods Ecol. Evol.*, 2041–210X.13272.

Bailey, M., Bush, S.R., Miller, A. & Kochen, M. (2016). The role of traceability in transforming seafood governance in the global South. *Curr. Opin. Environ. Sustain.*, 18, 25–32.

Baker, B. & Yeager, B. (2015). Coordinated Ocean Stewardship in the Arctic : Needs , Challenges and Possible Models for an Arctic Ocean Coordinating Agreement. *Transnatl. Environ. Law*, 4, 359–394.

Bal, P., Tulloch, A.I., Addison, P.F., McDonald-Madden, E. & Rhodes, J.R. (2018). Selecting indicator species for biodiversity management. *Front. Ecol. Environ.*, 16, 589–598.

Balmford, A., Bennun, L. & Brink, B. Ten. (2005). The Convention on Biological Diversity's 2010 target. *Science (80-.).*, 3, 212–213.

Ban, N.C., Mills, M., Tam, J., Hicks, C.C., Klain, S., Stoeckl, N., Bottrill, M.C., Levine, J., Pressey, R.L., Satterfield, T. & Chan, K.M. (2013). A social–ecological approach to conservation planning: embedding social considerations. *Front. Ecol. Environ.*, 11, 194–

202.

Baraldi, A.N. & Enders, C.K. (2010). An introduction to modern missing data analyses. *J. Sch. Psychol.*, 48, 5–37.

Barnes, M.D., Glew, L., Wyborn, C. & Craigie, I.D. (2018). Prevent perverse outcomes from global protected area policy. *Nat. Ecol. Evol.*, 2, 759–762.

Bartlett, L.J., Newbold, T., Purves, D.W., Tittensor, D.P. & Harfoot, M.B.J. (2016). Synergistic impacts of habitat loss and fragmentation on model ecosystems. *Proc. R. Soc. B Biol. Sci.*, 283.

Becker, W., Saisana, M., Paruolo, P. & Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecol. Indic.*, 80, 12–22.

Benini, L., Bandini, V., Marazza, D. & Contin, A. (2010). Assessment of land use changes through an indicator-based approach: A case study from the Lamone river basin in Northern Italy. *Ecol. Indic.*, 10, 4–14.

Bennett, J.R., Shaw, J.D., Terauds, A., Smol, J.P., Aerts, R., Bergstrom, D.M., Blais, J.M., Cheung, W.W., Chown, S.L., Lea, M.-A., Nielsen, U.N., Pauly, D., Reimer, K.J., Riddle, M.J., Snape, I., Stark, J.S., Tulloch, V.J. & Possingham, H.P. (2015). Polar lessons learned: long-term management based on shared threats in Arctic and Antarctic environments. *Front. Ecol. Environ.*, 13, 316–324.

Best, A., Blobel, D., Cavalieri, S., Giljum, S., Hammer, M., Lutter, S., Simmons, C. & Lewis, K. (2008). Potential of the Ecological Footprint for monitoring environmental impacts from natural resource use. *Sustain. Dev.*, 312.

Biggs, D., Cooney, R., Roe, D., Dublin, H.T., Allan, J.R., Challender, D.W.S. & Skinner, D. (2017). Developing a theory of change for a community-based response to illegal wildlife trade. *Conserv. Biol.*, 31, 5–12.

Biodiversity Indicators Partnership. (2018). Biodiversity Indicators Partnership [WWW Document]. *2018*. URL https://www.bipindicators.net/

Bird, K.J., Charpentier, R.R., Gautier, D.L., Houseknecht, D.W., Klett, T.R., Pitman, J.K., Moore, T.E., Schenk, C.J., Tennyson, M.E. & Wandrey, C.J. (2008). Circum-Arctic Resource Appraisal: Estimates of Undiscovered Oil and Gas North of the Arctic Circle. *Arctic*, 2000.

Bockstaller, C. & Girardin, P. (2003). How to validate environmental indicators. *Agric. Syst.*, 76, 639–653.

Böhringer, C. & Jochem, P.E.P. (2007). Measuring the immeasurable - A survey of sustainability indices. *Ecol. Econ.*, 63, 1–8.

Borja, Á., Elliott, M., Andersen, J.H., Berg, T., Carstensen, J., Halpern, B.S., Heiskanen, A.-S., Korpinen, S., Lowndes, J.S.S., Martin, G. & Rodriguez-Ezpeleta, N. (2016). Overview of Integrative Assessment of Marine Systems: The Ecosystem Approach in Practice. *Front. Mar. Sci.*, 3, 68.

Bradshaw, G.A. & Borchers, J.G. (2000). Uncertainty as Information : Narrowing the Science-policy Gap, 4, 1–14.

Branch, T. a, Watson, R., Fulton, E. a, Jennings, S., McGilliard, C.R., Pablico, G.T., Ricard, D. & Tracey, S.R. (2010). The trophic fingerprint of marine fisheries. *Nature*, 468, 431–435.

Brang, P., Courbaud, B., Fischer, A., Kissling-Naf, I., Pettenella, D., Schonenberger, W., Spork, J. & Grimm, V. (2002). Developing indicators for the sustainable management of mountain forests using a modelling approach. *For. Policy Econ.*, 4, 113–123.

ten Brink, B. (2006). *A Long-Term Biodiversity , Ecosystem and Awareness Research Network. Indicators as communication tools: an evolution towards composite indicators*. *ALTER-Net*.

Buckland, S.T., Magurran, a E., Green, R.E. & Fewster, R.M. (2005). Monitoring change in biodiversity through composite indices. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 360, 243–254.

Buhl-Mortensen, L., Buhl-Mortensen, P., Dolan, M.F.J. & Holte, B. (2015). The MAREANO programme – A full coverage mapping of the Norwegian off-shore benthic environment and fauna. *Mar. Biol. Res.*, 11, 4–17.

Buhl-Mortensen, L., Ellingsen, K.E., Buhl-Mortensen, P., Skaar, K.L. & Gonzalez-Mirelis, G. (2016). Trawling disturbance on megabenthos and sediment in the Barents Sea: chronic effects on density, diversity, and composition. *ICES J. Mar. Sci. J. du Cons.*, 73, i98–i114.

Buma, B. & Wessman, C. a. (2011). Disturbance interactions can impact resilience mechanisms of forests. *Ecosphere*, 2, art64.

Bunnefeld, N., Hoshino, E. & Milner-Gulland, E.J. (2011). Management strategy evaluation: A powerful tool for conservation? *Trends Ecol. Evol.*, 26, 441–447.

Burck, J. & Bals, C. (2011). *The Climate Change Performance Index Background and Methodology*. *Clim. Chang. Perform. Index*.

Burgass, M.J., Halpern, B.S., Nicholson, E. & Milner-Gulland, E.J. (2017). Navigating uncertainty in environmental composite indicators. *Ecol. Indic.*, 75, 268–278.

Burgman, M., Carr, A., Godden, L., Gregory, R., McBride, M., Flander, L. & Maguire, L. (2011a). Redefining expertise and improving ecological judgment. *Conserv. Lett.*, 4, 81–87.

Burgman, M.A., McBride, M., Ashton, R., Speirs-Bridge, A., Flander, L., Wintle, B., Fidler, F., Rumpff, L. & Twardy, C. (2011b). Expert status and performance. *PLoS One*, 6, 1–7.

Butchart, S.H.M., Di Marco, M. & Watson, J.E.M. (2016). Formulating Smart Commitments on Biodiversity: Lessons from the Aichi Targets. *Conserv. Lett.*, 9, 457–468.

Butchart, S.H.M., Stattersfield, A.J., Baillie, J., Bennun, L.A., Stuart, S.N., Akcakaya, H.R., Hilton-Taylor, C. & Mace, G.M. (2005). Using Red List Indices to measure progress towards the 2010 target and beyond. *Philos. Trans. R. Soc. B Biol. Sci.*, 360, 255–268.

CAFF. (2010). *Arctic Biodiversity Trends 2010 – Selected indicators of change*. CAFF International Secretariat, Akureyri, Iceland.

CAFF. (2014). Marine Ecosystem Monitoring [WWW Document]. URL http://www.caff.is/marine

CAFF & PAME. (2017). *Arctic protected areas: indicator report 2017*. Akureyi, Iceland.

Cash, D.W., Clark, W.C., Alcock, F., Dickson, N.M., Eckley, N., Guston, D.H., Ja, J. & Mitchell, R.B. (2003). Knowledge systems for sustainable development. *Proc. Natl. Acad. Sci.*

Cavalieri, D.J., Parkinson, C.L., Gloersen, P. & Zwally, H. (2015). Sea ice concentrations from Nimbus-7 SMMR and DMSP SMM/I-SSMIS passive microwave data. 1979-2015.

Certain, G., Skarpaas, O., Bjerke, J.W., Framstad, E., Lindholm, M., Nilsen, J.E., Norderhaug, A., Oug, E., Pedersen, H.C., Schartau, A.K., van der Meeren, G.I., Aslaksen, I., Engen, S., Garnåsjordet, P.A., Kvaløy, P., Lillegård, M., Yoccoz, N.G. & Nybø, S. (2011). The nature index: A general framework for synthesizing knowledge on the state of biodiversity. *PLoS One*, 6.

Chapin, F.S., Zavaleta, E.S., Eviner, V.T., Naylor, R.L., Vitousek, P.M., Reynolds, H.L., Hooper, D.U., Lavorel, S., Sala, O.E., Hobbie, S.E., Mack, M.C. & Díaz, S. (2000). Consequences of changing biodiversity. *Nature*, 405, 234–42.

Chatziparadeisis, A. (2007). Policy Impact of Composite Indicators. *Stat. Bundesamt*, 1–5.

Cheung, W.W.L., Lam, V.W.Y., Sarmiento, J.L., Kearney, K., Watson, R. & Pauly, D. (2009). Projecting global marine biodiversity impacts under climate change scenarios. *Fish Fish.*, 10, 235–251.

Chung, E.S. & Lee, K.S. (2009). Prioritization of water management for sustainability using hydrologic simulation model and multicriteria decision making techniques. *J. Environ. Manage.*, 90, 1502–1511.

Cinelli, M., Coles, S.R. & Kirwan, K. (2014). Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. *Ecol. Indic.*, 46, 138–148.

Coll, M., Shannon, L.J., Kleisner, K.M., Juan-Jordá, M.J., Bundy, A., Akoglu, A.G., Banaru, D., Boldt, J.L., Borges, M.F., Cook, A., Diallo, I., Fu, C., Fox, C., Gascuel, D., Gurney, L.J., Hattab, T., Heymans, J.J., Jouffre, D., Knight, B.R., Kucukavsar, S., Large, S.I., Lynam, C., Machias, A., Marshall, K.N., Masski, H., Ojaveer, H., Piroddi, C., Tam, J., Thiao, D., Thiaw, M., Torres, M.A., Travers-Trolet, M., Tsagarakis, K., Tuck, I., van der Meeren, G.I., Yemane, D., Zador, S.G. & Shin, Y.-J. (2016). Ecological indicators to capture the effects of fishing on biodiversity and conservation status of marine ecosystems. *Ecol. Indic.*, 60, 947–962.

Collen, B., Loh, J., Whitmee, S., Mcrae, L., Amin, R. & Baillie, J.E.M. (2009). Monitoring Change in Vertebrate Abundance: the Living Planet Index. *Conserv. Biol.*, 23, 317–327.

Collen, B. & Nicholson, E. (2014). Taking the measure of change. *Science (80-.).*, 346, 166–167.

Committee for Development Policy. (2018). *Voluntary National Review Reports– what do they report?*

Connors, B.M. & Cooper, A.B. (2014). Determining Decision Thresholds and Evaluating Indicators when Conservation Status is Measured as a Continuum. *Conserv. Biol.*, 28, 1626–1635.

Convention on Biological Diversity. (2006). *Framework for monitoring implementation of the achievement of the 2010 target and integration of targets into the thematic programmes of work*.

Convention on Biological Diversity. (2016). *XIII/28. Indicators for the Strategic Plan for Biodiversity 2011-2020 and the Aichi Biodiversity Targets*.

Convention on Biological Diversity. (2017). *Scenarios for the 2050 Vision for Biodiversity*. Montreal, Canada.

Convention on Biological Diversity. (2018a). *ANALYSIS OF THE CONTRIBUTION OF TARGETS ESTABLISHED BY PARTIES AND PROGRESS TOWARDS THE AICHI BIODIVERSITY TARGETS*. Montreal, Canada.

Convention on Biological Diversity. (2018b). *Analysis of the Contribution of Targets Established by Parties and Progress Towards the Aichi Biodiversity Targets*. Montreal, Canada.

Convention on Biological Diversity. (2019). *POST-2020 GLOBAL BIODIVERSITY FRAMEWORK: DISCUSSION PAPER: CBD/POST2020/PREP/1/1*.

Cooney, R., Roe, D., Dublin, H. and Booker, F. (2018). *Wild life, Wild Livelihoods: Involving Communities in Sustainable Wildlife Management and Combatting the Illegal Wildlife Trade*. Nairobi, Kenya.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R.G., Sutton, P. & van den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*.

Costanza, R., de Groot, R., Sutton, P., van der Ploeg, S., Anderson, S.J., Kubiszewski, I., Farber, S. & Turner, R.K. (2014). Changes in the global value of ecosystem services. *Glob. Environ. Chang.*, 26, 152–158.

Costelloe, B., Collen, B., Milner-Gulland, E.J., Craigie, I.D., Mcrae, L., Rondinini, C. & Nicholson, E. (2015). Global Biodiversity Indicators Reflect the Modeled Impacts of Protected Area Policy Change. *Conserv. Lett.*, 2020, 14–20.

Crépin, A.S., Gren, Å., Engström, G. & Ospina, D. (2017). Operationalising a social–ecological system perspective on the Arctic Ocean. *Ambio*, 46, 475–485.

Crotti, R. & Mashri, T. (2015). *The Travel & Tourism Competitiveness Report 2015. Asp. Tour. 9*. Geneva, Switzerland.

Daigle, R.M., Archambault, P., Halpern, B.S., Stewart Lowndes, J.S. & Côté, I.M. (2017). Incorporating public priorities in the Ocean Health Index: Canada as a case study. *PLoS One*, 12, e0178044.

Dale, V.H. & Beyeler, S.C. (2001). Challenges in the development and use of ecological indicators. *Ecol. Indic.*, 1, 3–10.

Dalpadado, P., Arrigo, K.R., Hjøllo, S.S., Rey, F., Ingvaldsen, R.B., Sperfeld, E., Van Dijken, G.L., Stige, L.C., Olsen, A. & Ottersen, G. (2014). Productivity in the Barents Sea - Response to recent climate variability. *PLoS One*, 9.

Davies, T.D. & Baum, J.K. (2012). Extinction Risk and Overfishing: Reconciling Conservation and Fisheries Perspectives on the Status of Marine Fishes. *Sci. Rep.*, 2, 1–9.

Deser, C., Walsh, J.E. & Timlin, M.S. (2000). Arctic sea ice variability in the context of recent atmospheric circulation trends. *J. Clim.*, 13, 617–633.

Desmet, P. & Cowling, R. (2004). Using the Species-Area Relationship to Set Baseline Targets for Conservation. *Ecol. Soc.*, 9, art11.

Devillers, R., Pressey, R.L., Grech, A., Kittinger, J.N., Edgar, G.J., Ward, T. & Watson, R. (2015). Reinventing residual reserves in the sea: are we favouring ease of establishment over need for protection? *Aquat. Conserv. Mar. Freshw. Ecosyst.*, 25, 480–504.

Dichmont, C.M. & Fulton, E.A. (2017). Fisheries Science and Participatory Management Strategy Evaluation: Eliciting Objectives, Visions and Systems Models. In: *Decis. Conserv. Nat. Resour. Manag.* (eds. Bunnefeld, N., Nicholson, E. & Milner-Gulland, E.J.). Cambridge University Press, Cambridge, UK, p. 275.

Dickson, R.R. & Brown, J. (1994). The production of North Atlantic Deep Water: Sources, rates, and pathways. *J. Geophys. Res.*, 99, 12319.

Doherty, T.S., Bland, L.M., Bryan, B.A., Neale, T., Nicholson, E., Ritchie, E.G. & Driscoll, D.A. (2018). Expanding the Role of Targets in Conservation Policy. *Trends Ecol. Evol.*, 33, 809–812.

Dorward, L.J., Mittermeier, J.C., Sandbrook, C. & Spooner, F. (2017). Pokémon Go: Benefits, Costs, and Lessons for the Conservation Movement. *Conserv. Lett.*, 10, 160–165.

Eamer, J., Donaldson, G.M., Gaston, a J., Kosobokova, K.N., Lárusson, K.F., Melnikov, I. a, Reist, J.D., Richardson, E., Staples, L. & von Quillfeldt, C.H. (2013). *Life linked to ice: A guide to sea-ice-associated biodiversity in this time of rapid change*. *CAFF Assess. Ser. No. 10*. Iceland.

Eamer, J., Russell, D.E., McRae, L., Böhm, M., Deinet, S., Collen, B. & Gill, M.J. (2012). *The Arctic Species Trend Index 2011*. Iceland.

Eea. (2010). *Assessing biodiversity in Europe — the 2010 report*. *Europe*.

Elfes, C.T., Longo, C., Halpern, B.S., Hardy, D., Scarborough, C., Best, B.D., Pinheiro, T. & Dutra, G.F. (2014). A regional-scale ocean health index for Brazil. *PLoS One*, 9, 1–11.

Elliott, M. (2014). Integrated marine science and management: Wading through the morass. *Mar. Pollut. Bull.*, 86, 1–4.

Elmgren, R., Blenckner, T. & Andersson, A. (2015). Baltic Sea management: Successes and failures. *Ambio*, 44, 335–344.

Engelhard, G.H., Righton, D.A. & Pinnegar, J.K. (2014). Climate change and fishing: A century of shifting distribution in North Sea cod. *Glob. Chang. Biol.*, 20, 2473–2483.

Erb, K.H., Haberl, H. & Plutzar, C. (2012). Dependency of global primary bioenergy crop potentials in 2050 on food systems, yields, biodiversity conservation and political stability. *Energy Policy*, 47, 260–269.

Eriksen, E., Rune, H., Gjøs, H. & Primicerio, R. (2017). Spatial and temporal changes in the Barents Sea pelagic compartment during the recent warming. *Prog. Oceanogr.*, 151, 206–226.

European Comission. (2010). 2010/477/EU: Commission decision of 1 September 2010 on criteria and methodological standards on good environmental status of marine waters. (notified under document C(2010) 5956). *Off. J. Eur. Comm.*

Evans, M.R., Bithell, M., Cornell, S.J., Dall, S.R.X., Diaz, S., Emmott, S., Ernande, B., Grimm, V., Hodgson, D.J., Lewis, S.L., Mace, G.M., Morecroft, M., Moustakas, A., Murphy, E., Newbold, T., Norris, K.J., Petchey, O., Smith, M., Travis, J.M.J. & Benton, T.G. (2013). Predictive systems ecology. *Proc. R. Soc. B Biol. Sci.*, 280, 20131452–20131452.

Failing, L. & Gregory, R. (2003). Ten common mistakes in designing biodiversity indicators for forest policy. *J. Environ. Manage.*, 68, 121–132.

Famurewa, S.M., Stenström, C., Asplund, M., Galar, D. & Kumar, U. (2014). Composite indicator for railway infrastructure management. *J. Mod. Transp.*, 22, 214–224.

FAO. (1995). *Code of Conduct for Responsible Fisheries*. Rome.

FAO. (2003). *Fisheries management. The ecosystem approach to fisheries. FAO Technical Guidelines for Responsible Fisheries, 4, Suppl. 2.*

FAO. (n.d.). Fisheries and aquaculture software. FishStatJ - software for fishery statistical time series. [WWW Document]. *FAO Fish. Aquac. Dep.*

Fauchald, P., Anker-nilssen, T., Barrett, R.T., Bust-, J.O., Bårdsen, B., Christensen-dalsgaard, S., Engen, S., Erikstad, K.E., Hanssen, S.A., Lorentsen, S., Moe, B., Reiertsen, T.K., Systad, G.H. & Publications, N. (2015). *The status and trends of seabirds breeding in Norway and Svalbard*.

Fay, G., Large, S.I., Link, J.S. & Gamble, R.J. (2013). Testing systemic fishing responses with ecosystem indicators. *Ecol. Modell.*, 265, 45–55.

Folke, C., Hahn, T., Olsson, P. & Norberg, J. (2005). Adaptive Governance of Social-Ecological Systems. *Annu. Rev. Environ. Resour.*, 30, 441–473.

Ford, J.D., McDowell, G. & Pearce, T. (2015). The adaptation challenge in the Arctic. *Nat. Clim. Chang.*, 5, 1046–1053.

Francis, T.B., Levin, P.S. & Harvey, C.J. (2011). The perils and promise of futures analysis in marine ecosystem-based management. *Mar. Policy*, 35, 675–681.

Di Franco, A., Marchini, A., Baiata, P., Milazzo, M. & Chemello, R. (2009). Developing a scuba trail vulnerability index (STVI): A case study from a Mediterranean MPA. *Biodivers. Conserv.*, 18, 1201–1217.

Franks, D.M., Davis, R., Bebbington, A.J., Ali, S.H., Kemp, D. & Scurrah, M. (2014). Conflict translates environmental and social risk into business costs. *Proc. Natl. Acad. Sci.*, 111, 7576–7581.

Frazier, M., Longo, C. & Halpern, B.S. (2016). Mapping uncertainty due to missing data in the global ocean health index. *PLoS One*, 11.

Freudenberg, M. (2003). Composite indicators of country performance: a critical assessment. *OECD Sci. Technol. Ind. Work. Pap.*, 16, 35.

Friedman, K., Garcia, S.M. & Rice, J. (2018). Mainstreaming biodiversity in fisheries. *Mar. Policy*, 95, 209–220.

Fulton, E. a., Smith, A.D.M. & Punt, A.E. (2005). Which ecological indicators can robustly detect effects of fishing? *ICES J. Mar. Sci.*, 62, 540–551.

Fulton, E.A. (2010). Approaches to end-to-end ecosystem models. *J. Mar. Syst.*, 81, 171–183.

Fulton, E.A., Link, J.S., Kaplan, I.C., Savina-Rolland, M., Johnson, P., Ainsworth, C., Horne, P., Gorton, R., Gamble, R.J., Smith, A.D.M. & Smith, D.C. (2011a). Lessons in modelling and management of marine ecosystems: the Atlantis experience. *Fish Fish.*, 12, 171–188.

Fulton, E.A., Punt, A.E., Dichmont, C.M., Harvey, C.J. & Gorton, R. (2018). Ecosystems say good management pays off. *Fish Fish.*, 1–31.

Fulton, E.A., Smith, A.D.M., Smith, D.C. & Johnson, P. (2014). An integrated approach is needed for ecosystem based fisheries management: Insights from ecosystem-level management strategy evaluation. *PLoS One*, 9.

Fulton, E.A., Smith, A.D.M., Smith, D.C. & Van Putten, I.E. (2011b). Human behaviour: The key source of uncertainty in fisheries management. *Fish Fish.*, 12, 2–17.

Game, E.T., Bremer, L.L., Calvache, A., Moreno, P.H., Vargas, A., Rivera, B. & Rodriguez, L.M. (2018). Fuzzy Models to Inform Social and Environmental Indicator Selection for Conservation Impact Monitoring. *Conserv. Lett.*, 11, 1–8.

Garciá Molinos, J., Halpern, B.S., Schoeman, D.S., Brown, C.J., Kiessling, W., Moore, P.J., Pandolfi, J.M., Poloczanska, E.S., Richardson, A.J. & Burrows, M.T. (2016). Climate velocity and the future global redistribution of marine biodiversity. *Nat. Clim. Chang.*, 6, 83–88.

Gari, S.R., Newton, A. & Icely, J.D. (2015). A review of the application and evolution of the DPSIR framework with an emphasis on coastal social-ecological systems. *Ocean Coast. Manag.*, 103, 63–77.

Garnett, S.T., Burgess, N.D., Fa, J.E., Fernández-Llamazares, Á., Molnár, Z., Robinson, C.J., Watson, J.E.M., Zander, K.K., Austin, B., Brondizio, E.S., Collier, N.F., Duncan, T., Ellis, E., Geyle, H., Jackson, M. V., Jonas, H., Malmer, P., McGowan, B., Sivongxay, A. & Leiper, I. (2018). A spatial overview of the global importance of Indigenous lands for conservation. *Nat. Sustain.*, 1, 369–374.

Gelman, A. & Hill, J. (2006). Missing-data imputation. In: *Data Anal. Using Regres. Multilevel/Hierarchical Model.* pp. 529–543.

Giampietro, M. & Saltelli, A. (2014). Footprints to nowhere. *Ecol. Indic.*, 46, 610–621.

Gisladottir, G. & Stocking, M. (2005). Land degradation control and its global environmental benefits. *L. Degrad. Dev.*, 16, 99–112.

von Glasow, R., Jickells, T.D., Baklanov, A., Carmichael, G.R., Church, T.M., Gallardo, L., Hughes, C., Kanakidou, M., Liss, P.S., Mee, L., Raine, R., Ramachandran, P., Ramesh, R., Sundseth, K., Tsunogai, U., Uematsu, M. & Zhu, T. (2013). Megacities and Large Urban Agglomerations in the Coastal Zone: Interactions Between Atmosphere, Land, and Marine Ecosystems. *Ambio*, 42, 13–28.

Government of Pakistan. (2017). *Pakistan National Biodiversity Strategy and Action Plan*.

Green, E.J., Buchanan, G.M., Butchart, S.H.M., Chandler, G.M., Burgess, N.D., Hill, S.L.L. & Gregory, R.D. (2019). Relating characteristics of global biodiversity targets to reported progress. *Conserv. Biol.*, cobi.13322.

Griffiths, C.L., Robinson, T.B., Lange, L. & Mead, A. (2010). Marine Biodiversity in South Africa: An Evaluation of Current States of Knowledge. *PLoS One*, 5, e12008.

Grønnevet, L. (2016). The joint Russian-Norwegian governance of the Barents Sea LME fisheries. *Environ. Dev.*, 17, 296–309.

Grupp, H. & Mogee, M.E. (2004). Indicators for national science and technology policy: How robust are composite indicators? *Res. Policy*, 33, 1373–1384.

Gullestad, P., Abotnes, A.M., Bakke, G., Skern-Mauritzen, M., Nedreaas, K. & Søvik, G. (2017). Towards ecosystem-based fisheries management in Norway – Practical tools for keeping track of relevant issues and prioritising management efforts. *Mar. Policy*, 77, 104–110.

Gullestad, P., Aglen, A., Bjordal, Å., Blom, G., Johansen, S., Krog, J., Misund, O.A. &

Røttingen, I. (2014). Changing attitudes 1970–2012: evolution of the Norwegian management framework to prevent overfishing and to secure long-term sustainability. *ICES J. Mar. Sci.*, 71, 173–182.

Hagerman, S. & Pelai, R. (2016). "As far as possible and as appropriate": Implementing the Aichi Biodiversity Targets. *Conserv. Lett.*, 53, 1689–1699.

Haine, T.W.N. & Martin, T. (2017). The Arctic-Subarctic sea ice system is entering a seasonal regime: Implications for future Arctic amplification. *Sci. Rep.*, 7, 4618.

Hák, T., Janoušková, S. & Moldan, B. (2016a). Sustainable Development Goals: A need for relevant indicators. *Ecol. Indic.*, 60, 565–573.

Hák, T., Janoušková, S. & Moldan, B. (2016b). Sustainable Development Goals: A need for relevant indicators. *Ecol. Indic.*, 60, 565–573.

Halpern, B.S., Frazier, M., Afflerbach, J., O'Hara, C., Katona, S., Stewart Lowndes, J.S., Jiang, N., Pacheco, E., Scarborough, C. & Polsenberg, J. (2017). Drivers and implications of change in global ocean health over the past five years. *PLoS One*, 12, 1–23.

Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., Lowndes, J.S., Rockwood, R.C., Selig, E.R., Selkoe, K. a. & Walbridge, S. (2015a). Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nat. Commun.*, 6, 7615.

Halpern, B.S., Frazier, M.R., Potapenko, J., Casey, K.S. & Koenig, K. (2015b). Cumulative human impacts: raw stressor data (2008 and 2013).

Halpern, B.S., Klein, C.J., Brown, C.J., Beger, M., Grantham, H.S., Mangubhai, S., Ruckelshaus, M., Tulloch, V.J., Watts, M., White, C. & Possingham, H.P. (2013a). Achieving the triple bottom line in the face of inherent trade-offs among social equity, economic return, and conservation. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 6229–6234.

Halpern, B.S., Longo, C., Hardy, D., McLeod, K.L., Samhouri, J.F., Katona, S.K., Kleisner, K., Lester, S.E., O'Leary, J., Ranelletti, M., Rosenberg, A. a, Scarborough, C., Selig, E.R., Best, B.D., Brumbaugh, D.R., Chapin, F.S., Crowder, L.B., Daly, K.L., Doney, S.C., Elfes, C., Fogarty, M.J., Gaines, S.D., Jacobsen, K.I., Karrer, L.B., Leslie, H.M., Neeley, E., Pauly, D., Polasky, S., Ris, B., St Martin, K., Stone, G.S., Sumaila, U.R. & Zeller, D.

(2012). An index to assess the health and benefits of the global ocean. *Nature*, 488, 615–20.

Halpern, B.S., Longo, C., Lowndes, J.S.S., Best, B.D., Frazier, M., Katona, S.K., Kleisner, K.M., Rosenberg, A. a., Scarborough, C. & Selig, E.R. (2015c). Patterns and Emerging Trends in Global Ocean Health. *PLoS One*, 10, e0117863.

Halpern, B.S., Longo, C., McLeod, K.L., Cooke, R., Fischhoff, B., Samhouri, J.F. & Scarborough, C. (2013b). Elicited preferences for components of ocean health in the California Current. *Mar. Policy*, 42, 68–73.

Halpern, B.S., Longo, C., Scarborough, C., Hardy, D., Best, B.D., Doney, S.C., Katona, S.K., McLeod, K.L., Rosenberg, A. a & Samhouri, J.F. (2014). Assessing the health of the U.S. west coast with a regional-scale application of the Ocean Health Index. *PLoS One*, 9, e98995.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C. V., Micheli, F., D'Agrosa, C., Fruno, J.F., Casey, K.S., Ebert, C., Fox, H.E., Fujita, R., Heinemann, D., Lenihan, H.S., Madin, E.M.P., Perry, M.T., Selig, E.R., Spalding, M., Steneck, R. & Watson, R. (2008). A Global Map of Human Impact on Marine Ecosystems, 319, 948–953.

Hamilton, L.C. & Lammers, R.B. (2011). Linking pan-Arctic human and physical data. *Polar Geogr.*, 34, 107–123.

Hammill, M.O., Stenson, G.B., Doniol-Valcroze, T. & Mosnier, A. (2015). Conservation of northwest Atlantic harp seals: Past success, future uncertainty? *Biol. Conserv.*, 192, 181–191.

Han, X., Smyth, R.L., Young, B.E., Brooks, T.M., De Lozada, A.S., Bubb, P., Butchart, S.H.M., Larsen, F.W., Hamilton, H., Hansen, M.C. & Turner, W.R. (2014). A biodiversity indicators dashboard: Addressing challenges to monitoring progress towards the Aichi biodiversity targets using disaggregated global data. *PLoS One*, 9.

Hansen, C., Drinkwater, K.F., Jähkel, A., Fulton, E.A., Gorton, R. & Skern-Mauritzen, M. (2019). Sensitivity of the Norwegian and Barents Sea Atlantis end-to-end ecosystem model to parameter perturbations of key species. *PLoS One*, 14, e0210419.

Hansen, C., Skern-Mauritzen, M., van der Meeren, G.I., Jahkel, A. & Drinkwater, K. (2016). *Set-up of the Nordic and Barents Seas (NoBa) Atlantis model*. *Fisk. og havet nr. 2*.

Harris, P.T., Macmillan-lawler, M., Kullerud, L. & Rice, J.C. (2017). Arctic marine conservation is not prepared for the coming melt. *ICES J. Mar. Sci.*, 1–11.

Hill, S.L.L., Harfoot, M., Purvis, A., Purves, D.W., Collen, B., Newbold, T., Burgess, N.D. & Mace, G.M. (2016). Reconciling Biodiversity Indicators to Guide Understanding and Action. *Conserv. Lett.*, 9, 405–412.

Hjermann, D.Ø., Ottersen, G. & Stenseth, N.C. (2004). Competition among fishermen and fish causes the collapse of Barents Sea capelin. *Proc. Natl. Acad. Sci. U. S. A.*, 101, 11679 LP – 11684.

Hoagland, P., Jacoby, J. & Schumacher, M.E. (2001). Law of the Sea. In: *Encycl. Ocean Sci.* Elsevier, pp. 432–443.

Hobday, A.J., Chambers, L.E. & Arnould, J.P.Y. (2015). Prioritizing climate change adaptation options for iconic marine species. *Biodivers. Conserv.*, 24, 3449–3468.

Hovelsrud, G.K., Mckenna, M. & Huntington, H.P. (2008). Marine mammal harvests and other interactions with humans. *Ecol. Appl.*, 18, 135–147.

Hovelsrud, G.K., Poppel, B., Van Oort, B. & Reist, J.D. (2011). Arctic societies, cultures, and peoples in a changing cryosphere. *Ambio*, 40, 100–110.

Howarth, L.M., Roberts, C.M., Thurstan, R.H. & Stewart, B.D. (2014). The unintended consequences of simplifying the sea: Making the case for complexity. *Fish Fish.*, 15, 690–711.

Hsu, A., Emerson, J., Johnson, L., Malik, O., Schwartz, J. & Jaiteh, M. (2014). *2014 Environmental Performance Index Full Report and Analysis*. New Haven, CT.

Hsu, A., Johnson, L. & LLoyd, A. (2013). *Measuring Progress: A Practical Guide from the Developers of the Environmental Performance Index*. New Haven.

Huntington, H.P., Quakenbush, L.T. & Nelson, M. (2016). Effects of changing sea ice on marine mammals and subsistence hunters in northern Alaska from traditional knowledge interviews. *Biol. Lett.*, 12, 20160198.

Hussey, N.E., Harcourt, R.G. & Auger-Methe, M. (2016). Now is the time to protect the Arctic. *Science (80-.).*, 354, 1243–1243.

ICES. (2018). *Cod (Gadus morhua) in subareas 1 and 2 (Northeast Arctic)*.

IndiSeas. (2019). A selection of indicators for evaluating and communicating ecosystem status [WWW Document]. URL http://www.indiseas.org/more-information

IOC/UNESCO. (2010). *Why monitor the Arctic Ocean?*

IPBES. (2016). *IPBES: The methodological assessment report on Scenarios and Models of Biodiversity and Ecosystem Services*. Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany.

IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland.

IUCN. (2017). IUCN Red List of Threatened Species. Version 2017-2. [WWW Document]. URL www.iucnredlist.org

IUCN & UNEP-WCMC. (2017). Protected Planet: The World Database on Protected Areas (WDPA)/The Global Database on Protected Areas Management Effectiveness (GD-PAME)] [On-line] (downloaded May 2017).

Jacobs, R., Smith, P. & Goddard, M. (2004). Measuring performance: An examination of composite performance indicators, 1–112.

Jago-on, K.A.B., Kaneko, S., Fujikura, R., Fujiwara, A., Imai, T., Matsumoto, T., Zhang, J., Tanikawa, H., Tanaka, K., Lee, B. & Taniguchi, M. (2009). Urbanization and subsurface environmental issues: An attempt at DPSIR model application in Asian cities. *Sci. Total Environ.*, 407, 3089–3104.

Jantke, K., Jones, K.R., Allan, J.R., Chauvenet, A.L.M., Watson, J.E.M. & Possingham, H.P.

(2018). Poor ecological representation by an expensive reserve system: Evaluating 35 years of marine protected area expansion. *Conserv. Lett.*, e12584.

Jenkins, C.N. & Joppa, L. (2009). Expansion of the global terrestrial protected area system. *Biol. Conserv.*, 142, 2166–2174.

Jones, J.P.G., Collen, B., Atkinson, G., Baxter, P.W.J., Bubb, P., Illian, J.B., Katzner, T.E., Keane, A., Loh, J., McDonald-Madden, E., Nicholson, E., Pereira, H.M., Possingham, H.P., Pullin, A.S., Rodrigues, A.S.L., Ruiz-Gutierrez, V., Sommerville, M. & Milner-Gulland, E.J. (2011). The why, what, and how of global biodiversity indicators beyond the 2010 target. *Conserv. Biol.*, 25, 450–7.

Jones, K.R., Klein, C.J., Halpern, B.S., Venter, O., Grantham, H., Kuempel, C.D., Shumway, N., Friedlander, A.M., Possingham, H.P. & Watson, J.E.M. (2018a). The Location and Protection Status of Earth's Diminishing Marine Wilderness. *Curr. Biol.*, 28, 2506-2512.e3.

Jones, K.R., Venter, O., Fuller, R.A., Allan, J.R., Maxwell, S.L., Negret, P.J. & Watson, J.E.M. (2018b). One-third of global protected land is under intense human pressure. *Science (80-.).*, 360, 788–791.

Jørgensen, S.E. (2008). Overview of the model types available for development of ecological models. *Ecol. Modell.*, 215, 3–9.

Jørgensen, S.E., Burkhard, B. & Müller, F. (2013). Twenty volumes of ecological indicators – An accounting short review. *Ecol. Indic.*, 28, 4–9.

van Katwijk, M.M., Thorhaug, A., Marbà, N., Orth, R.J., Duarte, C.M., Kendrick, G.A., Althuizen, I.H.J., Balestri, E., Bernard, G., Cambridge, M.L., Cunha, A., Durance, C., Giesen, W., Han, Q., Hosokawa, S., Kiswara, W., Komatsu, T., Lardicci, C., Lee, K.-S., Meinesz, A., Nakaoka, M., O'Brien, K.R., Paling, E.I., Pickerell, C., Ransijn, A.M.A. & Verduin, J.J. (2016). Global analysis of seagrass restoration: the importance of large-scale planting. *J. Appl. Ecol.*, 53, 567–578.

Kaufmann, D., Kraay, A. & Mastruzzi, M. (2011). The Worldwide Governance Indicators: Methodology and Analytical Issues. *Hague J. Rule Law*, 3, 220–246.

van de Kerk, G., Manuel, A. & Kleinjans, R. (2014). *Sustainable Society Index 2014*.

Van de Kerk, G. & Manuel, A.R. (2008). A comprehensive index for a sustainable society: The SSI - the Sustainable Society Index. *Ecol. Econ.*, 66, 228–242.

Kissling, W.D., Ahumada, J.A., Bowser, A., Fernandez, M., Fernández, N., García, E.A., Guralnick, R.P., Isaac, N.J.B., Kelling, S., Los, W., McRae, L., Mihoub, J.-B., Obst, M., Wee, B. & Hardisty, A.R. (2017). Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biol. Rev.*, in press.

Klein, C.J., Steinback, C., Watts, M., Scholz, A.J. & Possingham, H.P. (2010). Spatial marine zoning for fisheries and conservation. *Front. Ecol. Environ.*, 8, 349–353.

Klugman, J. (2010). *Human Development Report 2010 The Real Wealth of Nations : Pathways to Human Development*. *Hum. Dev.* UNDP, New York.

Koivurova, T. (2010). Limits and possibilities of the Arctic Council in a rapidly changing scene of Arctic governance. *Polar Rec. (Gr. Brit).*, 46, 146–156.

Kokko, H. (2005). Useful ways of being wrong. *J. Evol. Biol.*, 18, 1155–1157.

Kontrick, A. V. (2018). Microplastics and Human Health: Our Great Future to Think About Now. *J. Med. Toxicol.*, 14, 117–119.

Korzukhin, M.D., Ter-Mikaelian, M.T. & Wagner, R.G. (1996). Process versus empirical models: which approach for forest ecosystem management? *Can. J. For. Res.*, 26, 879–887.

Kovacs, K.M., Lydersen, C., Overland, J.E. & Moore, S.E. (2011). Impacts of changing sea-ice conditions on Arctic marine mammals. *Mar. Biodivers.*, 41, 181–194.

Kraska, J. (2011). *Arctic Security in an Age of Climate Change*. Cambridge University Press.

Kraxner, F., Nordström, E.-M., Havlík, P., Gusti, M., Mosnier, A., Frank, S., Valin, H., Fritz, S., Fuss, S., Kindermann, G., McCallum, I., Khabarov, N., Böttcher, H., See, L., Aoki, K., Schmid, E., Máthé, L. & Obersteiner, M. (2013). Global bioenergy scenarios – Future forest development, land-use implications, and trade-offs. *Biomass and Bioenergy*, 57, 86–96.

Laidler, G.J., Ford, J.D., Gough, W.A., Ikummaq, T., Gagnon, A.S., Kowal, S., Qrunnut, K. & Irngaut, C. (2009). Travelling and hunting in a changing Arctic: Assessing Inuit vulnerability to sea ice change in Igloolik, Nunavut. *Clim. Change*, 94, 363–397.

Laidre, K.L., Stern, H., Kovacs, K.M., Lowry, L., Moore, S.E., Regehr, E. V, Ferguson, S.H., Wiig, Ø., Boveng, P., Angliss, R.P., Born, E.W., Litovka, D., Quakenbush, L., Lydersen, C., Vongraven, D. & Ugarte, F. (2015). Arctic marine mammal population status, sea ice habitat loss, and conservation recommendations for the 21st century. *Conserv. Biol.*, 00, 1–14.

Lam, V.W.Y., Cheung, W.W.L. & Sumaila, U.R. (2016). Marine capture fisheries in the Arctic: Winners or losers under climate change and ocean acidification? *Fish Fish.*, 17, 335–357.

Large, S.I., Fay, G., Friedland, K.D. & Link, J.S. (2013). Defining trends and thresholds in responses of ecological indicators to fishing and environmental pressures. *ICES J. Mar. Sci.*, 70, 755–767.

Larrosa, C., Carrasco, L.R. & Milner-Gulland, E.J. (2016). Unintended Feedbacks: Challenges and Opportunities for Improving Conservation Effectiveness. *Conserv. Lett.*, 9, 316–326.

Larsen, J.N., Schweitzer, P. & Petrov, A. (2015). *Arctic Social Indicators*.

Latawiec, A.E., Strassburg, B.B., Brancalion, P.H., Rodrigues, R.R. & Gardner, T. (2015). Creating space for large-scale restoration in tropical agricultural landscapes. *Front. Ecol. Environ.*, 13, 211–218.

Law, E.A., Ferraro, P.J., Arcese, P., Bryan, B.A., Davis, K., Gordon, A., Holden, M.H., Iacona, G., Marcos Martinez, R., McAlpine, C.A., Rhodes, J.R., Sze, J.S. & Wilson, K.A. (2017). Projecting the performance of conservation interventions. *Biol. Conserv.*, 215, 142–151.

Leadley, P.W., Krug, C.B., Alkemade, R., Pereira, H.M., Sumaila, U.R., Walpole, M., Marques, A., Newbold, T., Teh, L.S.., van Kolck, J., Bellard, C., Januchowski-Hartley, S.R. & Mumby, P.J. (2013). *Progress towards the Aichi Biodiversity Targets : An assessment of biodiversity trends, policy scenarios and key actions (Global Biodiversity Outlook 4 (GBO-*

4) - Technical Report). *CBD Tech. Ser.*

Lebreton, L.C.M., van der Zwet, J., Damsteeg, J.-W., Slat, B., Andrady, A. & Reisser, J. (2017). River plastic emissions to the world's oceans. *Nat. Commun.*, 8, 15611.

Legg, C.J. & Nagy, L. (2006). Why most conservation monitoring is, but need not be, a waste of time. *J. Environ. Manage.*, 78, 194–199.

Lewis, E., MacSharry, B., Juffe-Bignoli, D., Harris, N., Burrows, G., Kingston, N. & Burgess, N.D. (2017). Dynamics in the global protected-area estate since 2004. *Conserv. Biol.*

Link, J.S., Yemane, D., Shannon, L.J., Coll, M., Shin, Y.J., Hill, L. & Borges, M.D.F. (2010). Relating marine ecosystem indicators to fishing and environmental drivers: an elucidation of contrasting responses. *ICES J. Mar. Sci.*, 67, 787–795.

Loeng, H. & Drinkwater, K. (2007). An overview of the ecosystems of the Barents and Norwegian Seas and their response to climate variability. *Deep Sea Res. Part II Top. Stud. Oceanogr.*, 54, 2478–2500.

Loh, J., Green, R.E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V. & Randers, J. (2005). The Living Planet Index: using species population time series to track trends in biodiversity. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 360, 289–95.

Longo, C.S., Frazier, M., Doney, S.C., Rheuban, J.E. & Halpern, B.S. (2017). Using the Ocean Health Index to identify opportunities and challenges to improving Southern Ocean ecosystem health. *Front. Mar. Sci.*, 4, 1–18.

Lowndes, J.S.S., Best, B.D., Scarborough, C., Afflerbach, J.C., Frazier, M.R., O'Hara, C.C., Jiang, N. & Halpern, B.S. (2017). Our path to better science in less time using open data science tools. *Nat. Ecol. Evol.*, 1, 0160.

Lowndes, J.S.S., Pacheco, E.J., Best, B.D., Scarborough, C., Longo, C., Katona, S.K. & Halpern, B.S. (2015). Best practices for assessing ocean health in multiple contexts using tailorable frameworks. *PeerJ*, 3, e1503.

M.J. Gill, K. Crane, R. Hindrum, P. Arneberg, I. Bysveen, N.V. Denisenko, V. Gofman, A. Grant-Friedman, G. Gudmundsson, R.R. Hopcroft, K. Iken, A. Labansen, O.S. Liubina, I.A. Melnikov, S.E. Moore, J.D. Reist, B.I. Sirenko, J. Stow, F. Ugarte, D. Vongrave, J.W. (2011). *Arctic Marine Biodiversity Monitoring Plan (CBMP-MARINE PLAN)*. CAFF International Secretariat, Akureyri, Iceland.

Mace, G.M. & Baillie, J.E.M. (2007). The 2010 biodiversity indicators: Challenges for science and policy. *Conserv. Biol.*, 21, 1406–1413.

Mace, G.M., Barrett, M., Burgess, N.D., Freeman, R., Cornell, S.E., Grooten, M. & Purvis, A. (2018). Aiming higher - bending the curve of biodiversity loss. *Nat. Sustain.*, 1, 448–451.

Mace, G.M., Collar, N.J., Gaston, K.J., Hilton-Taylor, C., Akçakaya, H.R., Leader-Williams, N., Milner-Gulland, E.J. & Stuart, S.N. (2008). Quantification of extinction risk: IUCN's system for classifying threatened species. *Conserv. Biol.*, 22, 1424–1442.

Mace, G.M., Reyers, B., Alkemade, R., Biggs, R., Chapin, F.S., Cornell, S.E., Díaz, S., Jennings, S., Leadley, P., Mumby, P.J., Purvis, A., Scholes, R.J., Seddon, A.W.R., Solan, M., Steffen, W. & Woodward, G. (2014). Approaches to defining a planetary boundary for biodiversity. *Glob. Environ. Chang.*, 28, 289–297.

Maher, P.T. (2007). Arctic tourism: a complex system of visitors, communities, and environments. *Polar Geogr.*, 30, 1–5.

Maher, P.T. (2017). Tourism Futures in the Arctic. In: *Interconnected Arct. — UArctic Congr. 2016*, Springer Polar Sciences (eds. Latola, K. & Savela, H.). Springer International Publishing, Cham, pp. 213–220.

Mäkelä, A., Río, M. del, Hynynen, J., Hawkins, M.J., Reyer, C., Soares, P., van Oijen, M. & Tomé, M. (2012). Using stand-scale forest models for estimating indicators of sustainable forest management. *For. Ecol. Manage.*, 285, 164–178.

Di Marco, M., Butchart, S.H.M., Visconti, P., Buchanan, G.M., Ficetola, G.F. & Rondinini, C. (2016). Synergies and trade-offs in achieving global biodiversity targets. *Conserv. Biol.*, 30, 189–195.

Margoluis, R., Stem, C., Salafsky, N. & Brown, M. (2009). Using conceptual models as a planning and evaluation tool in conservation. *Eval. Program Plann.*, 32, 138–147.

Marques, A., Pereira, H.M., Krug, C., Leadley, P.W., Visconti, P., Januchowski-Hartley, S.R., Krug, R.M., Alkemade, R., Bellard, C., Cheung, W.W.L., Christensen, V., Cooper, H.D., Hirsch, T., Hoft, R., van Kolck, J., Newbold, T., Noonan-Mooney, K., Regan, E.C., Rondinini, C., Sumaila, U.R., Teh, L.S.L. & Walpole, M. (2014). A framework to identify enabling and urgent actions for the 2020 Aichi Targets. *Basic Appl. Ecol.*, 15, 633–638.

Martell, S. & Froese, R. (2013). A simple method for estimating MSY from catch and resilience. *Fish Fish.*, 14, 504–514.

Mattoo, A. & Subramanian, A. (2012). Equity in Climate Change: An Analytical Review. *World Dev.*, 40, 1083–1097.

Maxwell, S.L., Fuller, R.A., Brooks, T.M. & Watson, J.E.M. (2016). Biodiversity: The ravages of guns, nets and bulldozers. *Nature*, 536, 143–145.

Maxwell, S.L., Milner-Gulland, E.J., Jones, J.P.G., Knight, A.T., Bunnefeld, N., Nuno, A., Bal, P., Earle, S., Watson, J.E.M. & Rhodes, J.R. (2015). Being smart about SMART environmental targets. *Science (80-.).*, 347, 1075–1076.

McBride, M.F., Garnett, S.T., Szabo, J.K., Burbidge, A.H., Butchart, S.H.M., Christidis, L., Dutson, G., Ford, H.A., Loyn, R.H., Watson, D.M. & Burgman, M.A. (2012). Structured elicitation of expert judgments for threatened species assessment: A case study on a continental scale using email. *Methods Ecol. Evol.*, 3, 906–920.

McCauley, D., Heffron, R., Pavlenko, M., Rehner, R. & Holmes, R. (2016). Energy justice in the Arctic: Implications for energy infrastructural development in the Arctic. *Energy Res. Soc. Sci.*, 16, 141–146.

McCauley, D.J., Pinsky, M.L., Palumbi, S.R., Estes, J. a., Joyce, F.H. & Warner, R.R. (2015). Marine defaunation: Animal loss in the global ocean. *Science (80-.).*, 347, 247–254.

McGinnis, M.D. & Ostrom, E. (2014). Social-ecological system framework: Initial changes and continuing challenges. *Ecol. Soc.*, 19, 30.

McGregor, V.L., Horn, P.L., Fulton, E.A. & Dunn, M.R. (2019). From data compilation to model validation: a comprehensive analysis of a full deep-sea ecosystem model of the Chatham Rise. *PeerJ*, 7, e6517.

McNie, E.C. (2007). Reconciling the supply of scientific information with user demands: an analysis of the problem and review of the literature. *Environ. Sci. Policy*, 10, 17–38.

Mcowen, C.J., Ivory, S., Dixon, M.J.R., Regan, E., Obrecht, A., Tittensor, D.P., Teller, A. & Chenery, A.M. (2016). Sufficiency and Suitability of Global Biodiversity Indicators for Monitoring Progress to 2020 Targets. *Conserv. Lett.*, 9, 489–494.

McQuatters-Gollop, A., Mitchell, I., Vina-Herbon, C., Bedford, J., Addison, P.F.E., Lynam, C.P., Geetha, P.N., Vermeulan, E.A., Smit, K., Bayley, D.T.I., Morris-Webb, E., Niner, H.J. & Otto, S.A. (2019). From Science to Evidence – How Biodiversity Indicators Can Be Used for Effective Marine Conservation Policy and Management. *Front. Mar. Sci.*, 6, 1–16.

McRae, L., Deinet, S. & Freeman, R. (2017). The diversity-weighted living planet index: Controlling for taxonomic bias in a global biodiversity indicator. *PLoS One*, 12, 1–20.

Mcrae, L., Deinet, S., Gill, M. & Collen, B. (2012). *The Arctic Species Trend Index. Tracking Trends in Arctic marine populations.* Iceland.

Mendola, D. & Volo, S. (2017). Building composite indicators in tourism studies: Measurements and applications in tourism destination competitiveness. *Tour. Manag.*, 59, 541–553.

Mendoza, G. a. & Prabhu, R. (2003). Fuzzy methods for assessing criteria and indicators of sustainable forest management. *Ecol. Indic.*, 3, 227–236.

Millennium Ecosystem Assessment. (2005). *Ecosystems and Human Well-being: Synthesis*. *Ecosystems*.

Milner-Gulland, E.J. (2018). Documenting and tackling the illegal wildlife trade: change and continuity over 40 years. *Oryx*, 52, 597–598.

Milner-Gulland, E.J., Garcia, S., Arlidge, W., Bull, J., Charles, A., Dagorn, L., Fordham, S., Graff Zivin, J., Hall, M., Shrader, J., Vestergaard, N., Wilcox, C. & Squires, D. (2018). Translating the terrestrial mitigation hierarchy to marine megafauna by-catch. *Fish Fish.*, 19, 547–561.

Miloslavich, P., Bax, N.J., Simmons, S.E., Klein, E., Appeltans, W., Aburto-Oropeza, O., Andersen Garcia, M., Batten, S.D., Benedetti-Cecchi, L., Checkley, D.M., Chiba, S., Duffy, J.E., Dunn, D.C., Fischer, A., Gunn, J., Kudela, R., Marsac, F., Muller-Karger, F.E., Obura, D. & Shin, Y.-J. (2018). Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. *Glob. Chang. Biol.*

Molenaar, E.J. (2015). The Oslo Declaration on High Seas Fishing in the Central Arctic Ocean. In: *L. Heininen, H. Exner-Pirot, J. Plouffe* (eds. L. Heininen, H.E.-P. & Plouffe, J.). Northern Research Forum, Akureyri, pp. 427–431.

Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B. & Worm, B. (2011). How Many Species Are There on Earth and in the Ocean? *PLoS Biol.*, 9, e1001127.

Morar, N., Toadvine, T. & Bohannan, B.J.M. (2015). Biodiversity at Twenty-Five Years: Revolution Or Red Herring? *Ethics, Policy Environ.*, 18, 16–29.

Moriarty, P.E., Hodgson, E.E., Froehlich, H.E., Hennessey, S.M., Marshall, K.N., Oken, K.L., Siple, M.C., Ko, S., Koehn, L.E., Pierce, B.D. & Stawitz, C.C. (2018). The need for validation of ecological indices. *Ecol. Indic.*, 84, 546–552.

Munda, G. (2005). "Measuring Sustainability": A Multi-Criterion Framework. *Environ. Dev. Sustain.*, 7, 117–134.

Munda, G. (2008). *Social Multi-Criteria Evaluation for a Sustainable Economy.* Springer-Verlag, Berlin Heidelberg.

Munda, G. & Nardo, M. (2009). Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. *Appl. Econ.*, 41, 1513–1523.

Munda, G., Nardo, M., Saisana, M. & Srebotnjak, T. (2009). Measuring uncertainties in composite indicators of sustainability. *Int. J. Environ. Technol. Manag.*, 11, 7.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD Publishing.

Newbold, T., Hudson, L.N., Hill, S.L.L., Contu, S., Lysenko, I., Senior, R.A., Börger, L., Bennett, D.J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M.J., Feldman, A., Garon, M., Harrison, M.L.K., Alhusseini, T., Ingram, D.J., Itescu, Y., Kattge, J., Kemp, V., Kirkpatrick, L., Kleyer, M., Correia, D.L.P., Martin, C.D., Meiri, S., Novosolov, M., Pan, Y., Phillips, H.R.P., Purves, D.W., Robinson, A., Simpson, J., Tuck, S.L., Weiher, E., White, H.J., Ewers, R.M., Mace, G.M., Scharlemann, J.P.W. & Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, 520, 45.

Newton, A.C. (2011). Implications of Goodhart's Law for monitoring global biodiversity loss. *Conserv. Lett.*, 4, 264–268.

Nicholson, E., Collen, B., Barausse, A., Blanchard, J.L., Costelloe, B.T., Sullivan, K.M.E., Underwood, F.M., Burn, R.W., Fritz, S., Jones, J.P.G., McRae, L., Possingham, H.P. & Milner-Gulland, E.J. (2012). Making robust policy decisions using global biodiversity indicators. *PLoS One*, 7, e41128.

Nicholson, E., Fulton, E.A., Brooks, T.M., Blanchard, R., Leadley, P., Metzger, J.P., Mokany, K., Stevenson, S., Wintle, B.A., Woolley, S.N.C., Barnes, M., Watson, J.E.M. & Ferrier, S. (2019). Scenarios and Models to Support Global Conservation Targets. *Trends Ecol. Evol.*, 34, 57–68.

Niemeijer, D. & de Groot, R.S. (2008a). A conceptual framework for selecting environmental indicator sets. *Ecol. Indic.*, 8, 14–25.

Niemeijer, D. & de Groot, R.S. (2008b). Framing environmental indicators: moving from causal chains to causal networks. *Environ. Dev. Sustain.*, 10, 89–106.

Northern Sea Route Information Office. (2013). Final statistics figures for transit navigation on the NSR in 2013 [WWW Document]. *North. Sea Route Inf. Off.* URL http://www.arctic-lio.com/node/209

Norway Ministry of Climate and Environment. (2016). *Norway's national biodiversity action plan*.

Norwegian Government. (2010). *Act relating to the management of wild living marine resources (Marine Resources Act)*.

Norwegian Government. (2017). *New Growth , Proud History. The Norwegian Government's Ocean Strategy*.

Norwegian Government. (2018). *The Clearing-House Mechanism of the Convention on Biological Diversity - 6th National Report*.

Norwegian ministry of Climate and Environment. (2016). Update of the integrated management plan for the Barents Sea – Lofoten area including an update of the delimitation of the marginal ice zone. *Meld.St. 20 Rep. to Storting (white Pap.*, 20, 56.

Ocean Health Index. (2015). The Ocean Health Index Assessment Manual [WWW Document]. URL http://ohi-science.org/manual/index.html

Ocean Health Index. (2016). ohicore version: v2016.1 core Ocean Health Index functions, [downloaded 08/12/2016].

Ochoa-Gaona, S., Kampichler, C., de Jong, B.H.J., Hernndez, S., Geissen, V. & Huerta, E. (2010). A multi-criterion index for the evaluation of local tropical forest conditions in Mexico. *For. Ecol. Manage.*, 260, 618–627.

OECD. (2003). *OECD Environmental Indicators: development, measurement and use*.

Ojima, D.S., Galvin, K. a & Turner II, B.L. (1994). The global impact of land-use change. to understand global change, natural scientists must consider the social context influencing human impact on environment. *Bioscience*, 44, 300–304.

Olsen, E., Fay, G., Gaichas, S., Gamble, R., Lucey, S. & Link, J.S. (2016). Ecosystem model skill assessment. Yes We Can! *PLoS One*, 11, 1–24.

Olsen, E., Kaplan, I.C., Ainsworth, C., Fay, G., Gaichas, S., Gamble, R., Girardin, R., Eide, C.H., Ihde, T.F., Morzaria-Luna, H.N., Johnson, K.F., Savina-Rolland, M., Townsend, H., Weijerman, M., Fulton, E.A. & Link, J.S. (2018). Ocean Futures Under Ocean Acidification, Marine Protection, and Changing Fishing Pressures Explored Using a Worldwide Suite of Ecosystem Models. *Front. Mar. Sci.*, 5, 1–23.

Onarheim, I.H. & Årthun, M. (2017). Toward an ice-free Barents Sea. *Geophys. Res. Lett.*, 44, 8387–8395.

Ortega-Cisneros, K., Cochrane, K. & Fulton, E.A. (2017). An Atlantis model of the southern Benguela upwelling system: Validation, sensitivity analysis and insights into ecosystem functioning. *Ecol. Modell.*, 355, 49–63.

Ostrom, E. (2009). A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science (80-.).*, 325, 419–422.

Van Oudenhoven, A.P.E., Petz, K., Alkemade, R., Hein, L. & De Groot, R.S. (2012). Framework for systematic indicator selection to assess effects of land management on ecosystem services. *Ecol. Indic.*, 21, 110–122.

Overland, J.E. & Wang, M. (2013). When will the summer Arctic be nearly sea ice free? *Geophys. Res. Lett.*, 40, 2097–2101.

Parker, C.F. & Karlsson, C. (2018). The UN climate change negotiations and the role of the United States: assessing American leadership from Copenhagen to Paris. *Env. Polit.*, 27, 519–540.

Parmentier, F.-J.W., Christensen, T.R., Sørensen, L.L., Rysgaard, S., McGuire, a. D., Miller, P. a. & Walker, D. a. (2013). The impact of lower sea-ice extent on Arctic greenhouse-gas exchange. *Nat. Clim. Chang.*, 3, 195–202.

Parr, T.W., Jongman, R.H.G. & Kulvik, M. (2010). *The selection of biodiversity indicators for EBONE*.

Paruolo, P., Saisana, M. & Saltelli, A. (2013). Ratings and rankings: Voodoo or science? *J. R. Stat. Soc. Ser. A Stat. Soc.*, 176, 609–634.

Patrício, J., Elliott, M., Mazik, K., Papadopoulou, K.-N. & Smith, C.J. (2016). DPSIR—Two Decades of Trying to Develop a Unifying Framework for Marine Environmental Management? *Front. Mar. Sci.*, 3.

Pauly, D. & Watson, R. (2005). Background and interpretation of the “Marine Trophic Index” as a measure of biodiversity. *Philos. Trans. R. Soc. B Biol. Sci.*, 360, 415–423.

Pauly, D. & Zeller, D. (2016). Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. *Nat. Commun.*, 7, 10244.

Pedersen, B., Nybø, S. & Sæther, S.A. (2016). *Nature Index for Norway 2015. Ecological framework, computational methods, database and information systems – NINA Report 1226.*

Van Pelt, T.I., Huntington, H.P., Romanenko, O.V. & Mueter, F.J. (2017). The missing middle: Central Arctic Ocean gaps in fishery research and science coordination. *Mar. Policy*, 85, 79–86.

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Höft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M. & Wegmann, M. (2013). Essential Biodiversity Variables. *Science (80-.).*, 339, 277 LP – 278.

Petter Johnsen, J. & Eliasen, S. (2011). Solving complex fisheries management problems: What the EU can learn from the Nordic experiences of reduction of discards. *Mar. Policy*, 35, 130–139.

Phelps, J., Webb, E.L. & Agrawal, A. (2010). Does REDD + Threaten to recentralize forest governance? *Science (80-.).*

Pinsky, M.L., Reygondeau, G., Caddell, R., Palacios-Abrantes, J., Spijkers, J. & Cheung, W.W.L. (2018). Preparing ocean governance for species on the move. *Science (80-.).*, 360, 1189–1191.

Pitcher, T., Kalikoski, D., Pramod, G. & Short, K. (2009a). Not honouring the code. *Nature*, 457, 658–659.

Pitcher, T.J., Kalikoski, D., Short, K., Varkey, D. & Pramod, G. (2009b). An evaluation of progress in implementing ecosystem-based management of fisheries in 33 countries. *Mar. Policy*, 33, 223–232.

Plagányi, É.E., Punt, A.E., Hillary, R., Morello, E.B., Thébaud, O., Hutton, T., Pillans, R.D., Thorson, J.T., Fulton, E.A., Smith, A.D.M., Smith, F., Bayliss, P., Haywood, M., Lyne, V. & Rothlisberg, P.C. (2014). Multispecies fisheries management and conservation: Tactical applications using models of intermediate complexity. *Fish Fish.*, 15, 1–22.

Poore, J. & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science (80-.).*, 360, 987 LP – 992.

Protected Planet. (2019). Norway Protected Area Coverage [WWW Document]. URL https://www.protectedplanet.net/country/NO

Protection of the Arctic Marine Environment. (2013). *The Arctic Ocean Review Project PHASE II Report, 2011-2013*.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A. & Haddon, M. (2016). Management strategy evaluation: best practices. *Fish Fish.*, 17, 303–334.

Purves, D.W., Scharlemann, J., Harfoot, M., Newbold, T., Tittensor, D.P., Hutton, J. & Emmott, S. (2013). Time to model all life on Earth. *Nature*, 493, 295–297.

Purvis, A., Newbold, T., De Palma, A., Contu, S., Hill, S.L.L., Sanchez-Ortiz, K., Phillips, H.R.P., Hudson, L.N., Lysenko, I., Börger, L. & Scharlemann, J.P.W. (2018). *Modelling and Projecting the Response of Local Terrestrial Biodiversity Worldwide to Land Use and Related Pressures: The PREDICTS Project*. *Adv. Ecol. Res.* 1st edn. Elsevier Ltd.

Rands, M.R.W., Adams, W.M., Bennun, L., Butchart, S.H.M., Clements, A., Coomes, D., Entwistle, A., Hodge, I., Kapos, V., Scharlemann, J.P.W., Sutherland, W.J. & Vira, B. (2010). Biodiversity conservation: challenges beyond 2010. *Science (80-.).*, 329, 1298–1303.

Rangel, T.F., Edwards, N.R., Holden, P.B., Diniz-Filho, J.A.F., Gosling, W.D., Coelho, M.T.P., Cassemiro, F.A.S., Rahbek, C. & Colwell, R.K. (2018). Modeling the ecology and evolution of biodiversity: Biogeographical cradles, museums, and graves. *Science (80-.).*, 361.

Regan, H., Colyvan, M. & Burgman, M. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecol. Appl.*, 12, 618–628.

Ricard, D., Minto, C., Jensen, O.P. & Baum, J.K. (2012). Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. *Fish Fish.*, 13, 380–398.

Rice, J.C. & Rochet, M.J. (2005). A framework for selecting a suite of indicators for fisheries management. *ICES J. Mar. Sci.*, 62, 516–527.

Rickels, W., Dovern, J., Hoffmann, J., Quaas, M.F., Schmidt, J.O. & Visbeck, M. (2016). Indicators for monitoring sustainable development goals: An application to oceanic development in the European Union. *Earth's Futur.*, 4, 252–267.

Riedler, B., Pernkopf, L., Strasser, T., Lang, S. & Smith, G. (2015). International Journal of Applied Earth Observation and Geoinformation A composite indicator for assessing habitat quality of riparian forests derived from Earth observation data. *Int. J. Appl. Earth Obs. Geoinf.*, 37, 114–123.

Rindorf, A., Mumford, J., Baranowski, P., Clausen, L.W., García, D., Hintzen, N.T., Kempf, A., Leach, A., Levontin, P., Mace, P., Mackinson, S., Maravelias, C., Prellezo, R., Quetglas, A., Tserpes, G., Voss, R. & Reid, D. (2017). Moving beyond the MSY concept to reflect multidimensional fisheries management objectives. *Mar. Policy*, 85, 33–41.

Ripple, W.J., Wolf, C., Newsome, T.M., Galetti, M., Alamgir, M., Crist, E., Mahmoud, M.I. & Laurance, W.F. (2017). World Scientists' Warning to Humanity: A Second Notice. *Bioscience*, 67, 1026–1028.

Robertson, D.P. & Hull, R.B. (2001). Beyond biology: Toward a more public ecology for conservation. *Conserv. Biol.*, 15, 970–979.

Rockström, J., Gaffney, O., Rogelj, J., Meinshausen, M., Nakicenovic, N. & Schellnhuber, H.J. (2017). A roadmap for rapid decarbonization. *Science (80-.).*, 355, 1269 LP – 1271.

Roll, U., Mittermeier, J.C., Diaz, G.I., Novosolov, M., Feldman, A., Itescu, Y., Meiri, S. & Grenyer, R. (2016). Using Wikipedia page views to explore the cultural importance of global reptiles. *Biol. Conserv.*, 204, 42–50.

Rosenberg, A.A., Kleisner, K.M., Afflerbach, J., Anderson, S.C., Dickey-Collas, M., Cooper, A.B., Fogarty, M.J., Fulton, E.A., Gutiérrez, N.L., Hyde, K.J.W., Jardim, E., Jensen, O.P., Kristiansen, T., Longo, C., Minte-Vera, C. V., Minto, C., Mosqueira, I., Osio, G.C., Ovando, D., Selig, E.R., Thorson, J.T., Walsh, J.C. & Ye, Y. (2017). Applying a New Ensemble Approach to Estimating Stock Status of Marine Fisheries Around the World. *Conserv. Lett.*, 00, 1–9.

Rowland, J.A., Nicholson, E., Murray, N.J., Keith, D.A., Lester, R.E. & Bland, L.M. (2018). Selecting and applying indicators of ecosystem collapse for risk assessments. *Conserv. Biol.*, 32, 1233–1245.

Sabatier, P.A. (1986). Top-Down and Bottom-Up Approaches to Implementation Research: A Critical Analysis and Suggested Synthesis. *J. Public Policy*.

Saha, A., McRae, L., Dodd, C.K., Gadsden, H., Hare, K.M., Lukoschek, V. & Böhm, M. (2018). Tracking Global Population Trends: Population Time-Series Data and a Living Planet Index for Reptiles. *J. Herpetol.*, 52, 259–268.

Saisana, M. & Philippas, D. (2012). *Sustainable Society Index (SSI): Taking societies' pulse along social, environmental and economic issues - The Joint Research Centre audit on the SSI*.

Saisana, M., Saltelli, a. & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *J. R. Stat. Soc. Ser. A (Statistics Soc.*, 168, 307–323.

Saisana, M. & Saltelli, A. (2011). Rankings and Ratings: Instructions for Use. *Hague J. Rule Law*, 3, 247–268.

Salomon, A.K., Gaichas, S.K., Jensen, O.P., Agostini, V.N., Sloan, N. a., Rice, J., McClanahan, T.R., Ruckelshaus, M.H., Levin, P.S., Dulvy, N.K. & Babcock, E. a. (2011). Bridging the divide between fisheries and marine conservation science. *Bull. Mar. Sci.*, 87, 251–274.

Samhouri, J.F., Lester, S.E., Selig, E.R., Halpern, B.S., Fogarty, M.J., Longo, C. & McLeod, K.L. (2012). Sea sick? Setting targets to assess ocean health and ecosystem services. *Ecosphere*, 3, 1–18.

Samhouri, J.F., Levin, P.S. & Ainsworth, C.H. (2010). Identifying thresholds for ecosystem-based management. *PLoS One*, 5.

Samuelsen, A., Huse, G. & Hansen, C. (2009). Shelf recruitment of Calanus finmarchicus off the west coast of Norway: role of physical processes and timing of diapause termination. *Mar. Ecol. Prog. Ser.*, 386, 163–180.

Sarkki, S., Niemelä, J., Tinch, R., Jäppinen, J.P., Nummelin, M., Toivonen, H. & Von Weissenberg, M. (2016). Are national biodiversity strategies and action plans appropriate for building responsibilities for mainstreaming biodiversity across policy sectors? The case of Finland. *J. Environ. Plan. Manag.*, 59, 1377–1396.

Sato, K., Inoue, J. & Watanabe, M. (2014). Influence of the Gulf Stream on the Barents Sea ice retreat and Eurasian coldness during early winter. *Environ. Res. Lett.*, 9, 084009.

Schmidt, J.I., Aanesen, M., Klokov, K.B., Khrutschev, S. & Hausner, V.H. (2015). Demographic and economic disparities among Arctic regions. *Polar Geogr.*, 0513, 1–20.

Scholes, R.J. & Biggs, R. (2005). A biodiversity intactness index. *Nature*, 434, 45–49.

Secretariat of the Convention on Biological Diversity. (2016). *Progress in the implementation of the convention and the Strategic Plan for Biodiversity 2011-2020 and towards the achievement of the Aichi Biodiversity Targets*.

Segan, D.B., Murray, K.A. & Watson, J.E.M. (2016). A global assessment of current and future biodiversity vulnerability to habitat loss-climate change interactions. *Glob. Ecol. Conserv.*, 5, 12–21.

Selig, E.R., Frazier, M., O׳Leary, J.K., Jupiter, S.D., Halpern, B.S., Longo, C., Kleisner, K.L., Sivo, L. & Ranelletti, M. (2015). Measuring indicators of ocean health for an island nation: The ocean health index for Fiji. *Ecosyst. Serv.*, 16, 403–412.

Selkoe, K.A., Blenckner, T., Caldwell, M.R., Crowder, L.B., Erickson, A.L., Essington, T.E., Estes, J.A., Fujita, R.M., Halpern, B.S., Hunsicker, M.E., Kappel, C. V., Kelly, R.P., Kittinger, J.N., Levin, P.S., Lynham, J.M., Mach, M.E., Martone, R.G., Mease, L.A., Salomon, A.K., Samhouri, J.F., Scarborough, C., Stier, A.C., White, C. & Zedler, J. (2015). Principles for managing marine ecosystems prone to tipping points. *Ecosyst. Heal. Sustain.*, 1, art17.

Serpetti, N., Baudron, A.R., Burrows, M.T., Payne, B.L., Helaouët, P., Fernandes, P.G. & Heymans, J.J. (2017). Impact of ocean warming on sustainable fisheries management informs the Ecosystem Approach to Fisheries, 1–15.

Serrao-Neumann, S., Davidson, J.L., Baldwin, C.L., Dedekorkut-Howes, A., Ellison, J.C., Holbrook, N.J., Howes, M., Jacobson, C. & Morgan, E.A. (2016). Marine governance to avoid tipping points: Can we adapt the adaptability envelope? *Mar. Policy*, 65, 56–67.

Shchepetkin, A.F. & McWilliams, J.C. (2005). The regional oceanic modeling system (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Model.*, 9, 347–404.

Shepherd, E., Milner-Gulland, E.J., Knight, A.T., Ling, M.A., Darrah, S., Soesbergen, A. & Burgess, N.D. (2016). Status and Trends in Global Ecosystem Services and Natural Capital: Assessing Progress Toward Aichi Biodiversity Target 14. *Conserv. Lett.*, 9, 429–437.

Shin, Y.-J., Houle, J.E., Akoglu, E., Blanchard, J.L., Bundy, A., Coll, M., Demarcq, H., Fu, C., Fulton, E.A., Heymans, J.J., Salihoglu, B., Shannon, L., Sporcic, M. & Velez, L. (2018). The specificity of marine ecological indicators to fishing in the face of environmental change: A multi-model evaluation. *Ecol. Indic.*, 89, 317–326.

Shin, Y.J., Bundy, A., Shannon, L.J., Simier, M., Coll, M., Fulton, E.A., Link, J.S., Jouffre, D., Ojaveer, H., MacKinson, S., Heymans, J.J. & Raid, T. (2010a). Can simple be useful and reliable? Using ecological indicators to represent and compare the states of marine ecosystems. *ICES J. Mar. Sci.*, 67, 717–731.

Shin, Y.J. & Shannon, L.J. (2010). Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. 1. The IndiSeas project. *ICES J. Mar. Sci.*, 67, 686–691.

Shin, Y.J., Shannon, L.J., Bundy, A., Coll, M., Aydin, K., Bez, N., Blanchard, J.L., Borges, M. d. F., Diallo, I., Diaz, E., Heymans, J.J., Hill, L., Johannesen, E., Jouffre, D., Kifani, S., Labrosse, P., Link, J.S., Mackinson, S., Masski, H., Mollmann, C., Neira, S., Ojaveer, H., ould Mohammed Abdallahi, K., Perry, I., Thiao, D., Yemane, D. & Cury, P.M. (2010b). Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. 2. Setting the scene. *ICES J. Mar. Sci.*, 67, 692–716.

Singh, G.G., Cisneros-Montemayor, A.M., Swartz, W., Cheung, W., Guy, J.A., Kenny, T.A., McOwen, C.J., Asch, R., Geffert, J.L., Wabnitz, C.C.C., Sumaila, R., Hanich, Q. & Ota, Y. (2018). A rapid assessment of co-benefits and trade-offs among Sustainable Development Goals. *Mar. Policy*, 93, 223–231.

Singh, R.K., Murty, H.R., Gupta, S.K. & Dikshit, a. K. (2009). An overview of sustainability assessment methodologies. *Ecol. Indic.*, 9, 189–212.

Skern-Mauritzen, M., Ottersen, G., Handegard, N.O., Huse, G., Dingsør, G.E., Stenseth, N.C. & Kjesbu, O.S. (2016). Ecosystem processes are rarely included in tactical fisheries management. *Metroeconomica*, 67, 165–175.

Skogen, M.D., Budgell, W.P. & Rey, F. (2007). Interannual variability in Nordic seas primary production. *ICES J. Mar. Sci.*, 64, 889–898.

Van Der Sluijs, J.P., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J. & Risbey, J. (2005). Combining Quantitative and Qualitative Measures of Uncertainty in Model-Based Environmental Assessment: The NUSAP System. *Risk Anal.*, 25, 481–492.

van der Sluijs, J.P., van Est, R. & Riphagen, M. (2010). Beyond consensus: Reflections from a democratic perspective on the interaction between climate politics and science. *Curr. Opin. Environ. Sustain.*, 2, 409–415.

Van der Sluijs, J.P., Potting, J., Risbey, J., van Vuuren, D., de Vries, B., Beusen, A., Heuberger, P., Corral Quintana, S., Funtowicz, S. & Kloprogge, P. (2002). Uncertainty assessment of the IMAGE/TIMER B1 CO2 emissions scenario, using the NUSAP method. *Rep. Comm. by Netherlands Natl. Res. Progr. Glob. Air Pollut. Clim. Chang. RIVM, Bilthoven, Netherlands*, 225.

Smith, C.J., Papadopoulou, K.-N., Barnard, S., Mazik, K., Elliott, M., Patrício, J., Solaun, O., Little, S., Bhatia, N. & Borja, A. (2016). Managing the Marine Environment, Conceptual Models and Assessment Considerations for the European Marine Strategy Framework Directive. *Front. Mar. Sci.*, 3.

Smith, L.C. & Stephenson, S.R. (2013). New Trans-Arctic shipping routes navigable by midcentury. *Proc. Natl. Acad. Sci. U. S. A.*, 110, E1191-5.

Sodhi, N.S., Butler, R. & Raven, P.H. (2011). Bottom-up conservation. *Biotropica*.

Starfield, A.M. (1997). A Pragmatic Approach to Modeling for Wildlife Management. *J. Wildl. Manage.*, 61, 261.

Steffen, W. & Stafford Smith, M. (2013). Planetary boundaries, equity and global sustainability: Why wealthy countries could benefit from more equity. *Curr. Opin. Environ. Sustain.*, 5, 403–408.

Stephens, A., Lewis, E.D. & Reddy, S.M. (2018). *Inclusive Systemic Evaluation (ISE4GEMs): A new approach for the SDG era*. New York.

Stephenson, P., Brooks, T.M., Butchart, S.H., Fegraus, E., Geller, G.N., Hoft, R., Hutton, J., Kingston, N., Long, B. & McRae, L. (2017). Priorities for big biodiversity data. *Front. Ecol. Environ.*, 15, 124–125.

Stewart, E., Dawson, J. & Johnston, M. (2015). Risks and opportunities associated with change in the cruise tourism sector: community perspectives from Arctic Canada. *Polar J.*, 5, 403–427.

Stewart, E.J., Dawson, J. & Draper, D. (2011). Cruise Tourism and Residents in Arctic Canada: Development of a Resident Attitude Typology. *J. Hosp. Tour. Manag.*, 18, 95–106.

van Strien, A.J., Meyling, A.W.G., Herder, J.E., Hollander, H., Kalkman, V.J., Poot, M.J.M., Turnhout, S., van der Hoorn, B., van Strien-van Liempt, W.T.F.H., van Swaay, C.A.M., van Turnhout, C.A.M., Verweij, R.J.T. & Oerlemans, N.J. (2016). Modest recovery of biodiversity in a western European country: The Living Planet Index for the Netherlands. *Biol. Conserv.*, 200, 44–50.

Sutherland, W.J. & Burgman, M.A. (2015). Use Experts Wisely. *Nature*, 526, 7–8.

Tam, J.C., Link, J.S., Rossberg, A.G., Rogers, S.I., Levin, P.S., Rochet, M.J., Bundy, A., Belgrano, A., Libralato, S., Tomczak, M., Van De Wolfshaar, K., Pranovi, F., Gorokhova, E., Large, S.I., Niquil, N., Greenstreet, S.P.R., Druon, J.N., Lesutiene, J., Johansen, M., Preciado, I., Patricio, J., Palialexis, A., Tett, P., Johansen, G.O., Houle, J. & Rindorf, A. (2017). Towards ecosystem-based management: Identifying operational food-web indicators for marine ecosystems. *ICES J. Mar. Sci.*, 74, 2040–2052.

Tesar, C., Dubois, M. & Shestakov, A. (2016). Toward strategic, coherent, policy-relevant arctic science. *Science (80-.).*, 353, 1368–1370.

Tittensor, D.P., Eddy, T.D., Lotze, H.K., Galbraith, E.D., Cheung, W., Barange, M., Blanchard, J.L., Bopp, L., Bryndum-Buchholz, A., Büchner, M., Bulman, C., Carozza, D.A., Christensen, V., Coll, M., Dunne, J.P., Fernandes, J.A., Fulton, E.A., Hobday, A.J., Huber, V., Jennings, S., Jones, M., Lehodey, P., Link, J.S., MacKinson, S., Maury, O., Niiranen, S., Oliveros-Ramos, R., Roy, T., Schewe, J., Shin, Y.J., Silva, T., Stock, C.A., Steenbeek, J., Underwood, P.J., Volkholz, J., Watson, J.R. & Walker, N.D. (2018). A protocol for the intercomparison of marine fishery and ecosystem models: Fish-MIP v1.0. *Geosci. Model Dev.*, 11, 1421–1442.

Tittensor, D.P., Walpole, M., Hill, S.L.L., Boyce, D.G., Britten, G.L., Burgess, N.D., Butchart, S.H.M., Leadley, P.W., Regan, E.C., Alkemade, R., Baumung, R., Bellard, C., Bouwman, L., Bowles-Newark, N.J., Chenery, A.M., Cheung, W.W.L., Christensen, V., Cooper, H.D., Crowther, A.R., Dixon, M.J.R., Galli, A., Gaveau, V., Gregory, R.D., Gutierrez, N.L., Hirsch, T.L., Hoft, R., Januchowski-Hartley, S.R., Karmann, M., Krug, C.B., Leverington, F.J., Loh, J., Lojenga, R.K., Malsch, K., Marques, A., Morgan, D.H.W., Mumby, P.J., Newbold, T., Noonan-Mooney, K., Pagad, S.N., Parks, B.C., Pereira, H.M., Robertson, T., Rondinini, C., Santini, L., Scharlemann, J.P.W., Schindler, S., Sumaila, U.R., Teh, L.S.L., van Kolck, J., Visconti, P. & Ye, Y. (2014). A mid-term analysis of progress toward international biodiversity targets. *Science (80-.).*, 346, 241–244.

Trujillo, P. (2008). Using a mariculture sustainability index to rank the performance of 53 countries. In: *Perform. 53 Ctries. Manag. Mar. Resour.* (eds. Alder, J. & Pauly, D.). Fisheries Centre Research Reports Fisheries Centre, University of British Columbia, Vancouver, Canada.

Tulloch, V.J.D., Plagányi, É.E., Matear, R., Brown, C.J. & Richardson, A.J. (2018). Ecosystem modelling to quantify the impact of historical whaling on Southern Hemisphere baleen whales. *Fish Fish.*, 19, 117–137.

Turner, R.K. (2000). Integrating natural and socio-economic science in coastal management. *J. Mar. Syst.*, 25, 447–460.

Turnhout, E., Hisschemöller, M. & Eijsackers, H. (2007). Ecological indicators: Between the two fires of science and policy. *Ecol. Indic.*, 7, 215–228.

UN Environment. (2018). *Emissions Gap Report 2018*.

UNEP/GRID Arendal. (2006). *The Arendal Seminar on multilateral environmental agreements and their relevance to the Arctic Arendal Norway*. *Seminar*.

UNEP. (2006). *Report on the eighth meeting of the Conference of the Parties to the Convention on Biological Diversity, CBD*. Nairobi.

United Nations. (2015). *Transforming our World: The 2030 Agenda for Sustainable Development.*

United Nations General Assembly. (1982). *Convention on the Law of the Sea*.

United Nations Statistical Commission. (2017). *Resolution adopted by the General Assembly on Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development (A/RES/71/313), Annex.*

Venter, O., Fuller, R.A., Segan, D.B., Carwardine, J., Brooks, T., Butchart, S.H.M., Di Marco, M., Iwamura, T., Joseph, L., O'Grady, D., Possingham, H.P., Rondinini, C., Smith, R.J., Venter, M. & Watson, J.E.M. (2014). Targeting Global Protected Area Expansion for Imperiled Biodiversity. *PLoS Biol.*, 12.

Venter, O., Magrach, A., Outram, N., Klein, C.J., Possingham, H.P., Di Marco, M. & Watson, J.E.M. (2018). Bias in protected-area location and its effects on long-term aspirations of biodiversity conventions. *Conserv. Biol.*, 32, 127–134.

Viken, A. (2011). Tourism, research, and governance on Svalbard: a symbiotic relationship. *Polar Rec. (Gr. Brit).*, 47, 335–347.

Visconti, P., Bakkenes, M., Baisero, D., Brooks, T., Butchart, S.H.M., Joppa, L., Alkemade, R., Di Marco, M., Santini, L., Hoffmann, M., Maiorano, L., Pressey, R.L., Arponen, A., Boitani, L., Reside, A.E., van Vuuren, D.P. & Rondinini, C. (2016). Projecting Global Biodiversity Indicators under Future Development Scenarios. *Conserv. Lett.*, 9, 5–13.

Vugteveen, P., Rouwette, E., Stouten, H., van Katwijk, M.M. & Hanssen, L. (2015). Developing social-ecological system indicators using group model building. *Ocean Coast. Manag.*, 109, 29–39.

Walpole, M., Almond, R.E.A., Besancon, C., Butchart, S.H.M., Campbell-Lendrum, D., Carr, G.M., Collen, B., Collette, L., Davidson, N.C., Dulloo, E., Fazel, A.M., Galloway, J.N., Gill, M., Goverse, T., Hockings, M., Leaman, D.J., Morgan, D.H.W., Revenga, C., Rickwood, C.J., Schutyser, F., Simons, S., Stattersfield, A.J., Tyrrell, T.D., Vie, J.-C. & Zimsky, M. (2009). Tracking Progress Toward the 2010 Biodiversity Target and Beyond. *Science (80-.).*, 325, 1503–1504.

Wassmann, P., Duarte, C., Agustí, S. & Sejr, M.K. (2011). Footprints of climate change in the Arctic marine ecosystem. *Glob. Chang. Biol.*, 17, 1235–1249.

Watson, J.E.M., Venter, O., Lee, J., Jones, K.R., Robinson, J.G., Possingham, H.P. & Allan, J.R. (2018). Protect the last of the wild. *Nature*, 563, 27–30.

Watson, R., Kitchingman, A., Gelchu, A. & Pauly, D. (2004). Mapping global fisheries: sharpening our focus. *Fish Fish.*, 5, 168–177.

Watson, R.A. (2017). A database of global marine commercial, small-scale, illegal and unreported fisheries catch 1950-2014. *Sci. Data*, 4, 1–9.

Weiss, C.H. (1997). How Can Theory-Based Evaluation Make Greater Headway? *Eval. Rev.*, 21, 501–524.

Wood, K.A., Stillman, R.A. & Hilton, G.M. (2018). Conservation in a changing world needs predictive models. *Anim. Conserv.*, 21, 87–88.

Woodhouse, E., Homewood, K.M., Beauchamp, E., Clements, T., McCabe, J.T., Wilkie, D. & Milner-Gulland, E.J. (2015). Guiding principles for evaluating the impacts of conservation interventions on human well-being. *Philos. Trans. R. Soc. B Biol. Sci.*, 370, 20150103.

Zeller, D., Booth, S., Pakhomov, E., Swartz, W. & Pauly, D. (2011). Arctic fisheries catches in Russia, USA, and Canada: baselines for neglected ecosystems. *Polar Biol.*, 34, 955–973.

Zoological Society of London. (2018). rlpi.

9 APPENDIX 1

Table A1-1 A description of different uncertainties as defined by Regan et al. (2002)

Type of Uncertainty		**Description**
Epistemic Uncertainty	Measurement Error	Results from imperfections in measurement equipment and observational techniques – includes operator and instrument error.
	Systematic Error	Occurs as a result of bias in the measuring equipment or the sampling procedure e.g. erroneous calibration of measurement equipment or judgement of scientist to include/exclude data.
	Natural Variation	Natural variation in a complex system which is difficult to predict.
	Inherent Randomness	This occurs not because of our limited understanding of a system but because of its inherent randomness.
	Model Uncertainty	Arises through uncertainties in models we use; either because they cannot replicate the entire system and inaccuracies in the constructs we use to represent processes within the system.
	Subjective Judgement	This results from the interpretation of data, particularly when it is sparse and error prone.
Linguistic Uncertainty	Vagueness	Arises as our language permits borderline cases e.g. “endangered” is vague because species are not just endangered or not endangered; some are borderline.
	Context Dependence	Results from failing to give context to understand a term e.g. merely describing a population as “small” leaves the reader wondering in doubt as to small compared with what.
	Ambiguity	Arises from a word having more than one meaning and not being clear about which is intended.
	Indeterminacy of Theoretical Terms	Occurs when present theoretical terms have potential for future ambiguity as yet unknown. Or where there are no accepted definitions to theoretical terms.
	Underspecificity	Occurs when a statement is not specific enough to be desired. E.g. a qualitative rather than a quantitative answer.

Table A1-2 Sources and types of uncertainty found within composite indicators

Source of Uncertainty		Epistemic Uncertainty						Linguistic Uncertainty				
		Measurement Error	**Systematic Error**	**Natural Variation**	**Inherent Randomness**	**Model Uncertainty**	**Subjective Judgement**	**Vagueness**	**Context Dependence**	**Ambiguity**	**Indeterminacy in theoretical terms**	**Underspecificity**
Theoretical Framework	Definition of Concept, Structure and Subgroups					✓	✓	✓	✓	✓	✓	✓
Data	Data Quality	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Missing Data		✓			✓	✓					
	Data Selection	±	±	±	±	±	±	±	±	±	±	±
	Sub-indicator Construction		✓			✓	✓					
	Data Normalization		✓			✓	✓					
	Data Structure Exploration					±	±					
Index Construction	Aggregation of Indicators					✓	✓					
	Weighting of Indicators	±	±	±	±	±	±	±	±	±	±	±
Statistical Coherence / Robustness	Uncertainty / Sensitivity Analysis					%						
Post Development	Composite Indicator Communication							±	±	±	±	±

✓ = *Source of uncertainty*
% = *Quantification of existing uncertainty*
± = *Amplification or reduction of existing uncertainty*
Types of uncertainty as defined by Regan et al. (2002)

Table A1-3 A selection of techniques to deal with missing data. Although modern imputation techniques may be the most robust, they are not widely used.

Technique	**Potential Methods**	**Description**
Discard Data	Complete case analysis, pairwise deletion.	Involves discarding data sets which contain gaps. Now generally preferred to keep data and work with the uncertainties than to simply discard as more sophisticated filling techniques are available.
Single Imputation	Mean imputation, regression imputation, stochastic regression imputation, hot deck imputation.	Many techniques where the researcher "fills in" missing data with generated values. Regression, stochastic and hot deck preferred to mean imputation, which produces unlikely estimates. Standard errors of estimates are often too low because of substantial uncertainty about the missing values. Choosing a single imputation in essence pretends that the true value is known with certainty (Gelman & Hill 2006).
Modern Imputation	Multiple imputation, maximum likelihood estimation	Considered superior to single imputation or discard techniques as they retain a full data set (more powerful) and produce unbiased estimates.
Adapted from (Baraldi & Enders 2010)		

10 APPENDIX 2

Table A2-1: Arctic Ocean Health Index Regions

Region	**EEZ Size (km^2)**	**Country**	**Notes**
Arctic Alaska	493,030	USA	Designated fishery control zone
Nunavut	1,481,161	Canada	Arctic territory
Canadian Beaufort Sea	712,277		Large Ocean Management Area
Russian Arctic	4,350,002	Russia	Land administrative regions not relevant at sea – territorial waters controlled by central government
Svalbard	796,484	Norway	Separate marine regions and data reporting for each Norwegian area
Arctic Norway	941,869		
Jan Mayen	291,801		
West Greenland	954,391	Greenland	Regions divided by FAO regions, which includes southern tip of Greenland
East Greenland	1,297,894		

Table A2-2: Pressure Matrix

goal	Element name	Chemical pollution	Chemical pollution 3nm	Pathogen Pollution	Nutrient pollution	Nutrient pollution 3nm	Trash	Invasive species	Genetic escapes	Subtidal softbottom destruction	Intertidal destruction	Commercial high bycatch fishing	Commerical low bycatch fishing	Artisinal low bycatch fishing	Targeted harvest	Sea surface temp	acidification	Sea Level rise	World governance indicator
AO			1			1		1		1	1	2	1		2	3		2	1
CP	Seaice shoreline															3		3	1
CW			3	3		3	3												1
ECO	Commercial Fishing	2			1			1	1	2	1	3	1	1					1
ECO	Tourism	3		3		3	3									3		2	1
ECO	Ocean Transport															1		1	
ECO	Food Processing	2			3	3		1		2						2		1	
FIS		1			1			1	1	2	1	3	1	1		3			1
HAB	Seaice edge															3		1	1
HAB			2			2		1		3		3	1	1					1
ICO			2			1	1	1			1	2			2	3	1		1
LIV	Commercial Fishing	2			1			1	1	2	1	3	1	1		3			1
LIV	Tourism	3		3		3	3									3		2	1
LIV	Transportation & Shipping							1										1	1
LIV	Education															3			

LIV	**Food Processing**	**2**			**3**	**3**		**1**		**2**						**2**		**1**	
LSP			**2**			**2**	**3**	**1**			**3**					**3**		**1**	**1**
MAR		**2**				**3**										**3**		**1**	**1**
NP		**1**			**2**			**1**								**3**	**1**	**1**	**1**
SPP		**2**			**3**		**1**	**1**	**1**	**3**	**2**	**3**	**1**	**1**	**1**	**3**	**1**	**1**	**1**
TR			**3**	**3**		**3**	**3**									**3**		**2**	**1**

The rank weights used in the pressures matrix were determined by Halpern *et al*. 2012 (*Nature*) based on scientific literature and expert opinion (see Supplemental Table S28 of *Halpern et al. 2012*). Scores from 1-3 are given to rank the importance of each pressure. Only values of 2 or 3 require that a resilience layer be activated when calculating the goal scores.

Stressors that have no impact are left blank in the matrix rather than being assigned a rank of zero, which would affect the average score.

Table A2-3: Resilience Matrix

goal	element	po_water	hd_mpa_coast	hd_mpa_eez	hd_habitat	sp_alien_species	fp_mpa_coast	fp_mpa_eez	fp_habitat	fp_mora	fp_mora_artisanal	g_tourism	g_mariculture	g_msi_gov	g_cites	species_diversity_3nm	species_diversity_eez	wgi_all	li_gci	li_sector_evenness
CP	Sea ice shoreline																			
CW		x																x		
ECO																		x	x	
FIS				x	x			x	x	x	x						x	x		
HAB	Sea ice edge																			
HAB	Soft bottom	x		x	x			x	x	x		x	x				x	x		
ICO		x		x	x			x	x	x	x				x		x	x		

SPP		x		x	x			x	x	x	x	x	x		x			x		
LIV																		x	x	x
LSP		x			x													x		
MAR		x											x	x				x		
NP			x		x		x								x	x		x		
TR		x																x		

Resilience is included in OHI as the sum of the ecological factors and social initiatives (policies, laws, etc.) that can positively affect goal scores by reducing or eliminating pressures.

goal	element	po_water	hd_mpa_coast	hd_mpa_eez	hd_habitat	sp_alien_species	fp_mpa_coast	fp_mpa_eez	fp_habitat	fp_mora	fp_mora_artisanal	g_tourism	g_mariculture	g_msi_gov	g_cites	species_diversity_3nm	species_diversity_eez	wgi_all	li_gci	li_sector_evenness
AO			x		x		x		x	x						x		x		
CP	seaice_shoreline																			
CW		x																x		
ECO																		x	x	
FIS				x	x			x	x	x	x						x	x		
HAB	seaice_edge																			
HAB	soft_bottom	x		x	x			x	x	x		x	x				x	x		
ICO		x		x	x			x	x	x	x				x		x	x		
SPP		x		x	x			x	x	x	x	x	x		x			x		
LIV																		x	x	x

LSP		x			x													x		
MAR		x											x	x				x		
NP			x		x		x								x	x		x		
TR		x																x		

Table A2-4: Full list of data layers included within AOHI

Targets	layer	filename	fld_value	units	Description
Artisanal Needs	ao_tend	ao_trend_gl2016.csv	score	Trend	Shoreline sea ice, marine mammal extinction risk, sustainability of artisanal fish stocks combined trend
Artisanal Needs	ao_status	ao_status_arc2016.csv	Score	Score	Shoreline sea ice, marine mammal extinction risk, sustainability of artisanal fish stocks combined status
Clean Waters	cw_chemical_trend	cw_chemical_trend_arc2016.csv	trend	trend score	Trends in chemical pollution
Clean Waters	cw_nutrient_trend	cw_nutrient_trend_arc2016.csv	trend	trend score	Trends in fertilizer pollution as a proxy for nutrient pollution
Clean Waters	cw_pathogen_trend	cw_pathogen_trend_arc2016.csv	trend	trend score	Trends in access to improved sanitation as a proxy for pathogen pollution trend
Clean Waters	cw_trash_trend	cw_trash_trend_arc2016.csv	trend	trend score	Trends in plastic pollution
Fisheries	fis_b_bmsy_arc2016	fis_b_bmsy_arc2016.csv	bbmsy	B / B_msy	B/Bmsy estimates obtained from RAM legacy or using the catch-MSY method
Fisheries	fis_meancatch_arc2016	fis_meancatch_arc2016.csv	mean_catch	metric tons	Catch data for each Taxon/FAO/AOHI region averaged across years
Food Provision	fp_wildcaught_weight	fp_wildcaught_weight_arc2016.csv	w_fis	proportion	Fisheries weighting factor
Habitat / Coastal Protection	hab_extent	hab_extent_arc2016.csv	km2	km^2	Habitat extent
Habitat / Coastal Protection	hab_health	hab_health_arc2016.csv	health	proportion	Habitat health

Habitat / Coastal Protection	hab_trend	hab_trend_arc2016.csv	trend	trend score	Habitat health trend
Iconic Species	ico_spp_iucn_status	ico_spp_iucn_status_arc2016.csv	category	IUCN risk category	IUCN risk category
Livelihoods and Economies	le_gdp	le_gdp_arc2016.csv	usd	2010 USD	GDP/sector
Livelihoods and Economies	le_jobs_sector_year	le_jobs_sector_year_arc2016.csv	value	jobs	Jobs
Livelihoods and Economies	le_unemployment	le_unemployment_arc2016.csv	percent	percent unemployed	Unemployment
Livelihoods and Economies	le_wage_sector_year	le_wage_sector_year_arc2016.csv	usd	2010 USD	Wages
Livelihoods and Economies	le_workforcesize_adj	le_workforcesize_adj_arc2016.csv	jobs	jobs	Total size of workforce (employed + unemployed)
Livelihoods and Economies	le_sector_weight	le_sector_weight_gl2016.csv	weight	value	Jobs weighting
Livelihoods and Economies	le_popn	le_population_arc2016.csv	count	count	Total population by subregion
Protected Places	lsp_prot_area_offshore3nm	lsp_prot_area_offshore3nm_arc2016.csv	area_km2	km^2	Coastal marine protected areas offshore 3km
Mariculture	mar_coastalpopn_inland25km	mar_coastalpopn_inland25km_sc2014-raster.csv	popsum	people	Coastal population inland 25 kilometers
Mariculture	mar_harvest_species	mar_harvest_species_arc2016.csv	species	species name	Mariculture species harvested
Mariculture	mar_harvest_tonnes	mar_harvest_tonnes_arc2016.csv	tonnes	tons	Mariculture harvest
Mariculture	mar_sustainability_score	mar_sustainability_score_arc2016.csv	sust_coeff	sustainability	Mariculture sustainability score
Marine Mammal Harvest	np_harvest	np_harvest_arc2016.csv	score	score	Marine mammal harvest score
pressures	cc_acid	cc_acid_gl2016.csv	pressure_score	pressure score	Ocean acidification
pressures	cc_slr	cc_slr_arc2016.csv	pressure_score	pressure score	Sea level rise

pressures	cc_sst	cc_sst_arc2016.csv	pressure_score	pressure score	Sea surface temperature (SST) anomalies
pressures	cc_uv	cc_uv_arc2016.csv	pressure_score	pressure score	UV radiation
pressures	fp_art_lb	fp_art_lb_arc2016.csv	pressure_score	pressure score	Low bycatch caused by artisanal fishing
pressures	fp_com_hb	fp_com_hb_arc2016.csv	pressure_score	pressure score	High bycatch caused by commercial fishing
pressures	fp_com_lb	fp_com_lb_arc2016.csv	pressure_score	pressure score	Low bycatch caused by commercial fishing
pressures	fp_targetharvest	fp_targetharvest_gl2016.csv	score	pressure score	Targeted harvest of cetaceans and sea turtles
pressures	hd_intertidal	hd_intertidal_gl2016.csv	pressure_score	pressure score	Coastal population density as a proxy for intertidal habitat destruction
pressures	hd_subtidal_sb	hd_subtidal_sb_arc2016.csv	pressures.score	pressure score	Demersal destructive commercial fishing practices relative to soft-bottom habitat area as a proxy for soft bottom habitat destruction
pressures	po_chemicals	po_chemicals_arc2016.csv	pressure_score	pressure score	Chemical pollution
pressures	po_nutrients	po_nutrients_arc2016.csv	pressure_score	pressure score	Fertilizer pollution as a proxy for nutrient pollution
pressures	sp_alien	sp_alien_gl2016.csv	pressures.score	pressure score	Alien species
pressures	sp_genetic	sp_genetic_gl2016.csv	pressures.score	pressure score	Introduced species as a proxy for genetic escapes
pressures	ss_wgi	ss_wgi_gl2016.csv	score	pressure score	Weakness of governance indicated with the WGI
pressures CW	po_chemicals_3nm	po_chemicals_3nm_arc2016.csv	pressure_score	pressure score	Coastal chemical pollution within 3 nm offshore
pressures CW	po_nutrients_3nm	po_nutrients_3nm_arc2016.csv	pressure_score	pressure score	Coastal fertilizer pollution as a proxy for nutrient pollution within 3nm offshore
pressures CW	po_pathogens	po_pathogens_arc2016.csv	pressure_score	pressure score	Access to improved sanitation as a proxy for pathogen pollution
pressures CW	po_trash	po_trash_arc2016.csv	pressure_score	pressure score	Trash pollution

pressures resilience	element_wts_cp_km2_x_protection	element_wts_cp_km2_x_protection_arc2016.csv	extent_rank	extent*rank	Used to weight elements of coastal protection goal
pressures resilience	element_wts_hab_pres_abs	element_wts_hab_pres_abs_arc2016.csv	boolean	boolean	Used to weight elements of habitat goal
resilience	fp_habitat	fp_habitat_gl2016.csv	resilience score	resilience score	CBD survey: habitat
resilience	fp_mora	fp_mora_gl2016.csv	value	resilience score	CBD survey: fishing
resilience	fp_mora_artisanal	fp_mora_artisanal_gl2016.csv	value	resilience score	CBD survey: artisanal fishing
resilience	fp_mpa_coast	fp_mpa_coast_arc2016.csv	resilience.score	resilience score	Protected marine area
resilience	fp_mpa_eez	fp_mpa_eez_arc2016.csv	resilience.score	resilience score	Protected marine area in eez
resilience	g_cites	g_cites_gl2016.csv	resilience_score	resilience score	Resilience from commitment to CITES
resilience	g_mariculture	g_mariculture_gl2016.csv	resilience score	resilience score	CBD survey: mariculture
resilience	g_msi_gov	g_msi_gov_gl2016.csv	resilience score	resilience score	MSI sustainability and regulations
resilience	g_tourism	g_tourism_gl2016.csv	resilience score	resilience score	CBD survey: tourism
resilience	hd_habitat	hd_habitat_gl2016.csv	resilience score	resilience score	CBD survey: habitat
resilience	hd_mpa_coast	hd_mpa_coast_arc2016.csv	resilience.score	resilience score	Protected marine area
resilience	hd_mpa_eez	hd_mpa_eez_arc2016.csv	resilience.score	resilience score	Protected marine area in eez
resilience	li_gci	li_gci_gl2016.csv	score	resilience score	GCI: competitiveness in achieving sustained economic prosperity
resilience	li_sector_evenness	li_sector_evenness_gl2016.csv	resilience score	resilience score	Sector evenness as a measure of economic diversity
resilience	po_water	po_water_gl2016.csv	resilience score	resilience score	CBD survey: water
resilience	sp_alien_species	sp_alien_species_gl2016.csv	resilience score	resilience score	CBD survey: alien species
resilience	species_diversity_3nm	species_diversity_3nm_arc2016.csv	score	resilience score	Coastal ecological integrity
resilience	species_diversity_eez	species_diversity_eez_arc2016.csv	score	resilience score	Ocean ecological integrity
resilience	wgi_all	wgi_all_gl2016.csv	score	resilience score	Strength of governance indicated with the WGI

spatial	rgn_area	rgn_area_arc2016.csv	area_km2	km^2	Region area of total EEZ ocean
spatial	rgn_area_offshore3nm	rgn_area_offshore3nm_arc2016.csv	area_km2	km^2	Region area offshore 3nm
spatial	rgn_georegion_labels	rgn_georegion_labels_arc2016.csv	label	label	Georegion labels per region, at 3 georegion levels
spatial	rgn_georegions	rgn_georegions_arc2016.csv	georgn_id	georegion id	Georegion ids per region, at 3 georegion levels
spatial	rgn_global	rgn_global_arc2016.csv	label	label	regions used in global analysis for Nature 2012, subset of regions_labels by type=eez and not deleted or disputed
spatial	rgn_labels	rgn_labels_gl2016.csv	label	label	regions by type (eez, subocean, unclaimed) and label
Species	spp_status	spp_status_arc2016.csv	score	status score	Species lists and IUCN threat categories as a proxy for iconic species status
Species	spp_trend	spp_trend_arc2016.csv	score	trend score	Species lists and IUCN threat categories as a proxy for iconic species trend
Tourism	tr_jobs_pct_tourism	tr_jobs_pct_tourism_arc2016.csv	pct	percent	Percent direct employment in tourism
Tourism	tr_jobs_total	tr_jobs_total_arc2016.csv	employed	people	Total labor force
Tourism	tr_jobs_tourism	tr_jobs_tourism_arc2016.csv	jobs	jobs	Direct employment in tourism
Tourism	tr_sustainability	tr_sustainability_gl2016.csv	S_score	score	Sustainability index
Tourism	tr_travelwarnings	tr_travelwarnings_gl2016.csv	multiplier	score	Travel warnings
Tourism	tr_unemployment	tr_unemployment_arc2016.csv	percent	percent	Percent unemployment

Fisheries

As per the global Fisheries goal, we based the reference point for sustainable yield on an estimate of the ratio of the most recent (2014) population biomass (B) to the biomass that can deliver maximum sustainable yield (B_{MSY}) for each taxon (B/B_{MSY}), where B/B_{MSY}=1 is the highest score. Each species' status score (SS) was calculated as:

$$SS = \begin{cases} B/B_{MSY} & if \quad B/B_{MSY} < 0.95 \\ 1 & if \quad B/B_{MSY} \leq 0.95\ B/B_{MSY} \leq 1.05 \\ max\{1 - \alpha(B/B_{MSY} - 1.05,\ \beta)\} & if \quad B/B_{MSY} > 1.05 \end{cases}$$

(Eq.S1)

For species with B/B_{MSY}<0.95 (using a 5% buffer for uncertainty), status declines directly proportional to the ratio of B to B_{MSY}. For species in which B/B_{MSY}>1.05, status declines at rate α, where $\alpha = 0.5$ ensures that underharvested species are penalised for distance from B_{MSY} at half the rate of heavily harvested species, to a minimum score of 0.25, β. The underharvest penalty was removed for Arctic Alaska, which is closed to industrial fishing, in a precautionary approach. As such it would be unfair to penalize any underharvest which the region has taken a strategic decision not to fish in this area. However, removing this penalty has no effect on overall scores.

Annual B/B_{MSY} time series for species fished in FAO regions 18 (Arctic Sea), 21 (Northwest Atlantic) and 27 (Northeast Atlantic) from the RAM Legacy database were used when available (http://ramlegacy.org, Ricard et al., 2012). For species not assessed in the RAM Legacy database we estimated annual B/BMSY scores utilising the data-limited 'catch-MSY' model, which uses catch data to estimate MSY of fish stocks (Martell & Froese 2013). The model was applied to catch time series using the 'datalimited' package in R (Anderson et al., 2017). For stocks where B/BMSY could not be estimated, we assigned the mean B/BMSY of species in the same year and FAO region. This was either because the species were identified to a coarser taxonomic level (e.g., family or class level) or had inadequate data which caused model failure (often because time series were not long enough or catches not high enough).

Species status score was then multiplied by a taxonomic penalty, which penalises reporting at lower than the species level, with increasing penalty for coarser reporting, which is considered a sign of poor management (Table S1). Finally status is calculated as the mean of stock status scores weighted by average catch measured throughout the time series for each region.

Table A2-5: Taxonomic Reporting Multipliers

Reporting Level	Penalty
Species	1
Genus	0.9
Family	0.8
Order	0.6
Class	0.25
Other	0.1

Mariculture

Sustainability scores are calculated for each species in each country based three criteria (fishmeal use, waste treatment, and seed and larvae origin criteria) from the Mariculture Sustainability Index. These criteria represent the internal mariculture practices with the potential to affect the long term sustainability of the mariculture system. The MSI reports data for 359 country-species combinations (with 60 countries and 86 species represented) for each assessment criterion. Scores for each assessment criterion were aggregated and averaged based on the proportion of production that each assessed species contributed to the overall production in each country in the current year. Country average scores were then rescaled from 0 to 1 using the maximum possible raw MSI score of 10 and minimum of 1, and then weighted equally to calculate a composite resilience.

Table A2-6: Mariculture sustainability criteria (Trujillo 2008)

Criteria	Description of practice and score scheme
Fishmeal use	Fish protein and oil inclusion in the diet at any stage of development must be considered; herbivore species will score 10, and carnivorous (piscivorous) organisms will score closer to 1, depending on the level of feed supplied.
Waste treatment	Water exchange, output destinations, recycling and filtering of open water discharge or closed system reuse systems. Systems that are closed score high (10), while open systems without waste treatments score low (1)
Seed and larvae origin	Hatcheries are major providers of larvae, fry and seeds. Broodstock origin and strain will also affect the score. Wild seed collection and its importance contribute to a low score due to bycatch and other effects on non-target species.

Livelihoods and Economies

Table A2-7: Employment and Wage Data Sources

Region	Link
Arctic Alaska	http://live.laborstats.alaska.gov/qcew/ http://live.laborstats.alaska.gov/labforce/
Nunavut	http://www.stats.gov.nu.ca/en/Labour%20and%20employment.aspx http://www.stats.gov.nu.ca/en/Economic%20income.aspx
Canadian Beaufort	http://www.statsnwt.ca/labour-income/labour-force- http://www.statsnwt.ca/labour-income/income/index.html
Russian Arctic	http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/regional_statistics/ http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/wages/labour_costs/
Svalbard	https://www.ssb.no/
Arctic Norway	https://www.ssb.no/
Greenland	http://www.stat.gl/dialog/topmain.asp?lang=en&subject=Labour%20Market&sc=AR http://www.stat.gl/dialog/topmain.asp?lang=en&subject=Income&sc=IN

Table A2-8: Marine Sectors considered for Livelihoods and Economies

Region	Marine sector
Arctic Alaska	Tourism
	Transport
Nunavut	Fishing
	Tourism
	Transport
Canadian Beaufort	Tourism
	Transport
Arctic Russia	Fishing
	Tourism
	Transport
Svalbard	Education
	Tourism
	Transport
Arctic Norway	Fishing
	Seafood
	Tourism
	Transport

Table A2-9: Sources of revenue data for marine sectors

Region	Link
Arctic Alaska	http://live.laborstats.alaska.gov/qcew/
Nunavut	http://www.stats.gov.nu.ca/en/Economic%20GDP.aspx
Canadian Beaufort	http://www.statsnwt.ca/economy/gdp/
Russian Arctic	http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/accounts/#
Svalbard	https://www.ssb.no/
Arctic Norway	https://www.ssb.no/
Greenland	http://www.stat.gl/dialog/topmain.asp?lang=en&subject=National%20Accounts&sc=NR

Iconic Species

Table A2-10: Iconic Species for each region

Common Name	Scientific name	Region
Arctic char	Salvelinus alpinus	All
Blue whale	Balaenoptera musculus	All
Bowhead whale	Balaena mysticetus	All
Common guillemot	Uria aalge	All
Common Eider	Somateria mollissima	All
Fin whale	Balaenoptera physalus	All
Humpback Whale	Megaptera novaeangliae	All
Minke whale	Balaenoptera acutorostrata	All
Polar bear	Ursus maritimus	All
Red knot	Calidris canutus	All
Sei whale	Balaenoptera borealis	All
Thick-billed guillemot	Uria lomvia	All
Bearded seal	Erignathus barbatus	Arctic Alaska
Beluga Whale	Delphinapterus leucas	Arctic Alaska
Ringed seal	Pusa hispida	Arctic Alaska
Spectacled Eider	Somateria fischeri	Arctic Alaska
Spotted seal	Phoca largha	Arctic Alaska
Steller's Eider	Polysticta steller	Arctic Alaska
Walrus	Odobenus rosmarus	Arctic Alaska
Beluga Whale	Delphinapterus leucas	Nunavut
Harbour Porpoise	Phocoena phocoena	Nunavut
Harbour seal	Phoca vitulina	Nunavut
Hooded Seal	Cystophora cristata	Nunavut
Ivory Gull	Pagophila eburnea	Nunavut
Killer Whale	Orcinus orca	Nunavut
Narwhal	Monodon monoceros	Nunavut
Ringed seal	Pusa hispida	Nunavut
Walrus	Odobenus rosmarus	Nunavut
Beluga Whale	Delphinapterus leucas	Canadian Beaufort Sea
Harbour Porpoise	Phocoena phocoena	Canadian Beaufort Sea
Harbour seal	Phoca vitulina	Canadian Beaufort Sea
Hooded Seal	Cystophora cristata	Canadian Beaufort Sea
Ivory Gull	Pagophila eburnea	Canadian Beaufort Sea
Killer Whale	Orcinus orca	Canadian Beaufort Sea
Narwhal	Monodon monoceros	Canadian Beaufort Sea
Ringed seal	Pusa hispida	Canadian Beaufort Sea
Walrus	Odobenus rosmarus	Canadian Beaufort Sea
Harbour porpoise	Phocoena phocoena	Arctic Russia
Harbour seal	Phoca vitulina	Arctic Russia
Narwhale	Monodon monoceros	Arctic Russia
North Atlantic right whale	Eubalaena glacialis	Arctic Russia
Walrus	Odobenus rosmarus	Arctic Russia
White-beaked dolphin	Lagenorhynchus albirostris	Arctic Russia
Arctic skua	Stercorarius parasiticus	Svalbard
Arctic tern	Sterna paradisaea	Svalbard

Bearded seal	Erignathus barbatus	Svalbard
Black guillemot	Cepphus grylle	Svalbard
Fulmar	Fulmarus glacialis	Svalbard
Glaucous gull	Larus hyperboreus	Svalbard
Harbour seal	Phoca vitulina	Svalbard
Hooded seal	Cystophora cristata	Svalbard
Ivory gull	Pagophila eburnea	Svalbard
King eider	Somateria spectabilis	Svalbard
Little auk	Alle alle	Svalbard
Narwhal	Monodon monoceros	Svalbard
Pink-footed Goose	Anser brachyrhynchus	Svalbard
Pomarine skua	Stercorarius pomarinus	Svalbard
Purple sandpipers	Calidris maritima	Svalbard
Ringed plover	Charadrius hiaticula	Svalbard
Ringed seal	Pusa hispida	Svalbard
Sanderling	Calidris alba	Svalbard
Walrus	Odobenus rosmarus	Svalbard
White-beaked dolphin	Lagenorhynchus albirostris	Svalbard
Arctic skua	Stercorarius parasiticus	Arctic Norway
Arctic tern	Sterna paradisaea	Arctic Norway
Bearded seal	Erignathus barbatus	Arctic Norway
Black guillemot	Cepphus grylle	Arctic Norway
Salmon	Salmo salar	Arctic Norway
Arctic tern	Sterna paradisaea	Jan Mayen
Bearded seal	Erignathus barbatus	Jan Mayen
Arctic skua	Stercorarius parasiticus	West Greenland
Arctic tern	Sterna paradisaea	West Greenland
Bearded seal	Erignathus barbatus	West Greenland
Beluga whale	Delphinapterus leucas	West Greenland
Black guillemot	Cepphus grylle	West Greenland
Bottlenosed dolphin	Tursiops truncatus	West Greenland
Fulmar	Fulmarus glacialis	West Greenland
Glaucous gull	Larus hyperboreus	West Greenland
Harbour porpoise	Phocoena phocoena	West Greenland
Harbour seal	Phoca vitulina	West Greenland
Hooded seal	Cystophora cristata	West Greenland
Ivory gull	Pagophila eburnea	West Greenland
Killer whale	Orcinus orca	West Greenland
King Eider	Somateria spectabilis	West Greenland
Little auk	Alle alle	West Greenland
Narhwal	Monodon monoceros	West Greenland
North Atlantic right whale	Eubalaena glacialis	West Greenland
Pink-footed Goose	Anser brachyrhynchus	West Greenland
Pomarine	Stercorarius pomarinus	West Greenland
Purple sandpiper	Calidris maritima	West Greenland
Ringed Plover	Charadrius hiaticula	West Greenland

Ringed seal	Pusa hispida	West Greenland
Salmon	Salmo salar	West Greenland
Sanderling	Calidris alba	West Greenland
White-beaked dolphin	Lagenorhynchus albirostris	West Greenland
Arctic skua	Stercorarius parasiticus	East Greenland
Arctic tern	Sterna paradisaea	East Greenland
Bearded seal	Erignathus barbatus	East Greenland
Beluga whale	Delphinapterus leucas	East Greenland
Black guillemot	Cepphus grylle	East Greenland
Bottlenosed dolphin	Tursiops truncatus	East Greenland
Fulmar	Fulmarus glacialis	East Greenland
Glaucous gull	Larus hyperboreus	East Greenland
Harbour porpoise	Phocoena phocoena	East Greenland
Harbour seal	Phoca vitulina	East Greenland
Hooded seal	Cystophora cristata	East Greenland
Ivory gull	Pagophila eburnea	East Greenland
Killer whale	Orcinus orca	East Greenland
King Eider	Somateria spectabilis	East Greenland
Little auk	Alle alle	East Greenland
Narhwal	Monodon monoceros	East Greenland
North Atlantic right whale	Eubalaena glacialis	East Greenland
Pink-footed Goose	Anser brachyrhynchus	East Greenland
Pomarine	Stercorarius pomarinus	East Greenland
Purple sandpiper	Calidris maritima	East Greenland
Ringed Plover	Charadrius hiaticula	East Greenland
Ringed seal	Pusa hispida	East Greenland
Salmon	Salmo salar	East Greenland
Sanderling	Calidris alba	East Greenland
White-beaked dolphin	Lagenorhynchus albirostris	East Greenland

Table A2-11: Score conversions for species trend

Species Trend	Trend Score
Increasing	0.025
Decreasing	-0.025

Marine Mammal Harvest

Table A2-12: Harvested Marine Mammals included for each region

Region	Scientific Name	Common Name
Arctic Alaska	*Odobenus rosmarus*	Walrus
	Phoca largha	Spotted Seal
Nunavut	*Monodon monoceros*	Narwhal
Arctic Russia	*Odobenus rosmarus*	Walrus
	Pagophilus groenlandicu	Harp Seal
Norway	*Pagophilus groenlandicu*	Harp Seal
Jan Mayen	*Pagophilus groenlandicu*	Harp Seal
	Cystophora cristata	Hooded Seal
West Greenland	*Monodon monoceros*	Narwhal
	Odobenus rosmarus	Walrus
East Greenland	*Monodon monoceros*	Narwhal
	Odobenus rosmarus	Walrus

Equation S2 below is a representation of Figure 2 in the main text.

$$S' = \begin{cases} 2.1 - SS & \text{when SS} > 1.1 \\ 1 & \text{when } 0.9 \leq SS \leq 1.1 \\ 0.25 + \dfrac{0.75}{0.90} * SS & \text{when SS} < 0.9 \end{cases}$$

(Eq. S2)

The 0.25 indicates the lowest value that can be obtained when the catch is lower than the catch limit (i.e. under-harvest); this value was used because under-harvesting can be beneficial to rebuild populations. The 0.75/0.90 establishes the slope from the minimum value (0.25) to the lower buffer of the ideal score (1.0 – 10% buffer applied to account for uncertainty). The 2.1 establishes the slope for over-harvesting from the upper buffer limit (1.0+10% buffer applied for uncertainty) so that a score of zero is achieved when Catch is twice or more than twice the Catch Limit.

Habitats

Soft Bottomed Habitat: Global catch data from Sea Around Us Project is provided at 0.5 degree raster scale in units of tonnes/km2 for each species and fishing gear type. We identified catch from trawling gear, defined as dredges, hand dredges, bottom trawls, and shrimp trawls (mid-water trawls were excluded). We summed the trawled catch data for each AOHI region for each year, and converted to catch density by dividing the annual catch by the area of trawlable (soft-bottom) habitat. 'Trawlable habitat' within a region was defined as shallow subtidal (0-60m) and outer shelf (60-200m) soft bottom habitat. We rescaled AOHI regions based on the 95th percentile global log transformed value from all year-country possibilities. Condition was then calculated as one minus the rescaled catch density in the most recent year and further rescaled to the global median condition value across all years, and any value greater than the median was set = 1.0. This follows the current Global OHI approach.

Species

Table A2-13: Score conversions for IUCN Red List Species

IUCN Red List Status	Status Score
Least concern/Lower risk (LC)	0
Near threatened (NT)	0.2
Vulnerable (VU)	0.4
Endangered (EN)	0.6
Critically Endangered	0.8
Extinct	1
Data Deficient	NA

Tourism and Recreation

The Travel and Tourism Competitiveness Index is produced by the World Economic Forum and measures the factors and policies that make a country an attractive place to invest in the travel and tourism sector (WEF 2015, http://reports.weforum.org/travel-and-tourism-competitiveness-report-2015). The index analyzes 140 countries and scores each based on three sub-indices: human, cultural, and natural resources; business environment and infrastructure; and regulatory framework. These three sub-indices are in turn composed of 14 “pillars” of Travel & Tourism Competitiveness that are informed by a multitude of individual indicators based on the World Economic Forum’s annual Executive Opinion Survey and data from publically available sources: human, cultural, and natural resources (human resources, affinity for travel and tourism, natural resources, and cultural resources); business environment and infrastructure (air transport infrastructure, ground transport infrastructure, tourism infrastructure, ICT infrastructure, and price competitiveness in the industry); and regulatory framework (policy rules and regulations, environmental sustainability, safety and security, health and hygiene, and prioritization of travel and tourism). Because these indicators are meant to represent the overall quality and future potential of the tourism sector within a country, we assume they are representative of the long term sustainability of the tourism sector within each country. Values range from 1-6.

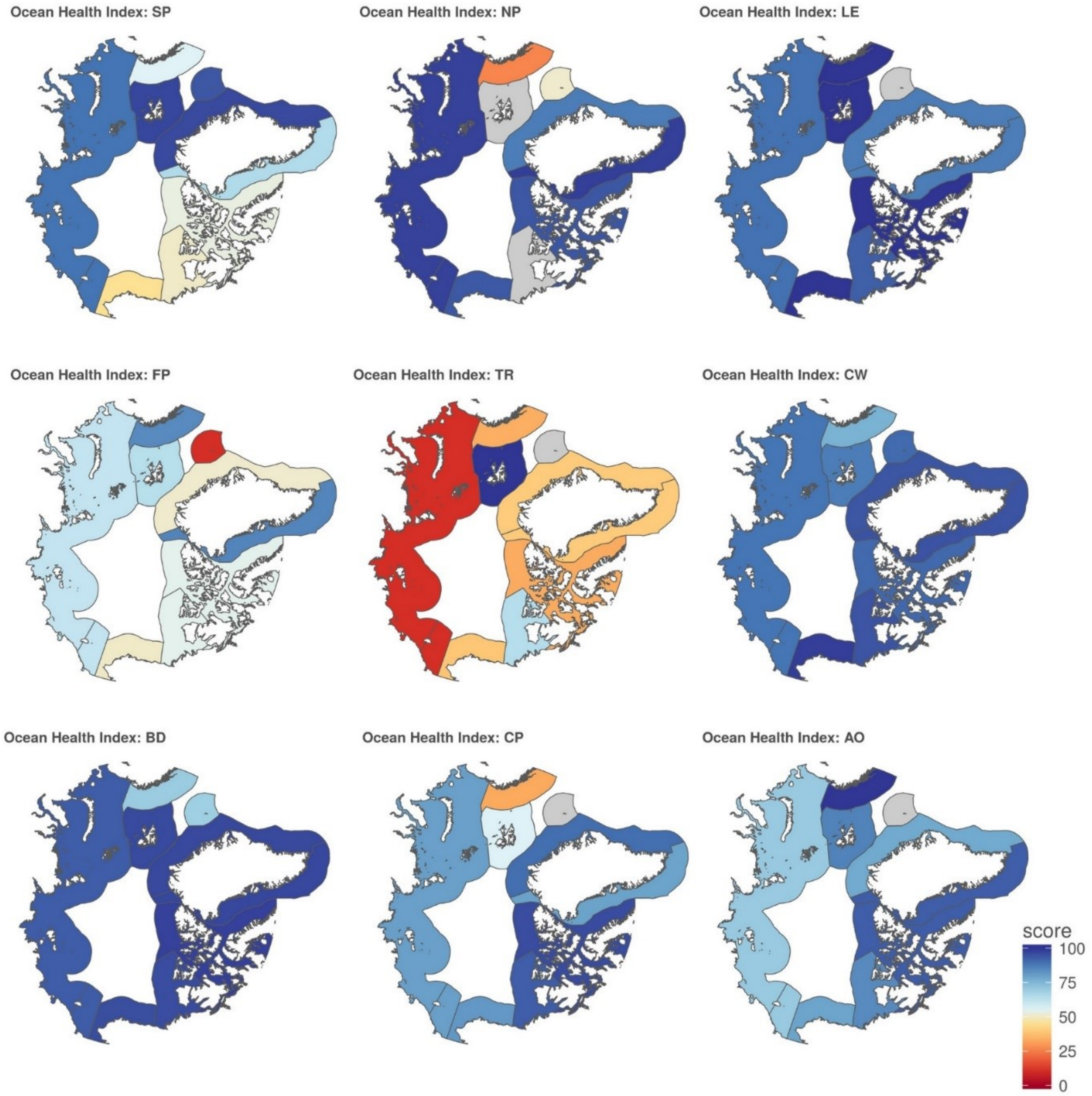

Figure A2-1: Maps of Arctic Ocean Health Index scores for each assessment region. Scores range from 0 (bad) to 100 (excellent). Grey areas indicate that particular goal was not relevant to that region and thus not assessed.

Table A2-14: Arctic Data portals

Region	Portal	Web address	Notes

Circumpolar	Circumpolar Arctic Coastal Communities Observatory Network (CACCON)	http://caccon.org/	Initiative aiming to build knowledge hubs to support, sustain and share adaptation for coastal communities. CACCON creates knowledge to support evidence-based decision making to adapt to climatic and socioeconomic changes
	ArcticData	http://arcticdata.is/	ArcticData provides access to data collected and developed through the activities of CAFF and PAME.
	Arctic Portal	http://arcticportal.org/	The Arctic Portal is a network of information and data sharing and serves as host to many web sites in a circumpolar context, supporting co-operation and outreach in science, education, and policy making. Includes an interactive mapping system, standardized permafrost monitoring data and the Arctic Transportation Database (log of ports and airports).
	ArcticStat	http://www.arcticstat.org/	ArcticStat is a permanent, public and independent statistical database dealing with the countries, regions and populations of the Circumpolar Arctic.
	Armap (Arctic Research Mapping Application)	http://armap.org/	ARMAP encompasses scientific research projects across the Arctic, funded or coordinated by multiple agencies and organizations. ARMAP uses best practices with information and mapping technologies to provide a comprehensive perspective in support of Arctic science.
	Arctic Data Explorer	http://nsidc.org/acadis/search/	Links to a selection of repositories
	Advanced Cooperative Arctic Data and Information Service (ACADIS)	https://www.aoncadis.org	Joint effort by the National Snow and Ice Data Center (NSIDC), the University Corporation for Atmospheric Research (UCAR), UNIDATA, and the National Center for Atmospheric Research (NCAR) to provide data archival, preservation and access for all projects funded by NSF's Arctic Science Program (ARC). ACADIS builds on the CADIS project that supported the Arctic Observing Network (AON). This portal will continue to be a gateway for AON data and is being expanded to include all NSF ARC data.
	Atlas of Community-Based Monitoring	http://www.arcticcbm.org	Provides a map of CBM projects in the Arctic and information about them including links to PIs or websites. It is intended to serve as an inventory of initiatives that will assist with network building and identification of best

			practices and challenges for the field. A secondary phase of the project will draw on CBM initiatives inventoried by the atlas, as well as a literature review, and interviews and input from practitioners, to draft a review of the state of CBM in the Arctic.
	Arctic Biodiversity Data Service (ABDS)	http://www.abds.is/	The ABDS is the data-management framework for managing data generated via CAFF and its Circumpolar Biodiversity Monitoring Programme (CBMP). It is an online, interoperable data management system which will serve as a focal point and common platform for all CAFF programs and projects as well as be a dynamic source for up-to-date circumpolar Arctic biodiversity information and emerging trends.
	Arctic Environmental Atlas	http://maps.grida.no/arctic/	Interactive atlas of the Arctic. Produced by the GRID-Arendal Centre, a centre collaborating with the United Nations Environment Programme (UNEP).
Canada	Nodicana D	http://www.cen.ulaval.ca/nordicanad	Environmental data from monitoring sites all over Canada.
	Arctic Science and Technology Information System (ASTIS)	http://www.aina.ucalgary.ca/astis/	The Arctic Science and Technology Information System (ASTIS) database contains 81,000 records describing publications and research projects about northern Canada.
Norway	Nordregio	http://www.nordregio.se/en/	Nordregio's main areas of research include regional development - urban and rural, city regional planning, demography, governance and gender, innovation and green growth, and sustainable development in the Arctic. Research competencies include the production of high-quality maps, the web-mapping tool NordMap and the development of state of the art statistical databases.
	ArcticWeb	http://www.arcticweb.com/	ArcticWeb exists to simplify access to public data sources in the Arctic Region. The information (via search and map interfaces) is used by oil and service companies for the purpose of exploration, early field development, environmental risk analysis, emergency preparedness, safety assessments and more. ArcticWeb covers the entire Norwegian Continental Shelf with data from a wide-range of Norwegian key data owners.

11 APPENDIX 3

Table A3-1: Components of the NoBa Atlantis Model

Full name	Species Included
Polar bear	
Killer whale	
Sperm whale	
Humpback whale	
Minke whale	
Fin whale	
Bearded seal	
Harp seal	
Hooded seal	
Ringed seal	
Arctic sea birds[2]	
Boreal sea birds	
Sharks, other	Picked dogish, Porbeagle, Tope shark
Demersals, other	Ling, Tusk
Pelagic large	Atlantic salmon
Pelagic small	Lumpfish, Norway pout
Redfish, other	Golden redfish
Demersal, large	Monkfish, Atlantic halibut, Atlantic wolfish, northern wolfish, spotted wolfish
Flatfish, other	European plaice, common dab, winter flounder
Long rough dab	
Skates and rays	Arctic skate, starry ray, sailray, longnosed skate, thornback ray, round skate, spinytail skate
Mesopelagic fish	Silvery lightfish, glacier lantern fish
Greenland halibut	
Mackerel	
Haddock	
Saithe	
Redfish	
Blue whiting	
Norwegian spring spawning herring	
Northeast arctic cod	
Polar cod	
Capelin	
Prawn	Pandalus borealis
Cephalopods	Gonatus fabricii
Red king crab	
Snow crab	
Gelatineous zooplankton	Aurelia aurita, cyanea capilate
Large zooplankton	Thysanoessa inermis
Medium zooplankton	Parameterized as Calanus finmarchicus
Small zooplankton	Small copepods, oncaea, pseudocalanus, (Oithona similis)
Dinoflagellates	
Small phytoplankton	Flagellates
Large phytoplankton	Diatoms

Predatory benthos	Echinoderms, sea urchins, annelids and anemones
Detrivore benthos	Selected annelids, echinoderms
Benthic filter feeders	Selected molluscs, barnacles, moss animals, anemones (Tridonta borealis)
Sponges	Geodia baretti
Corals	Lophelia pertusa
Pelagic bacteria	
Benthic bacteria	
Refractory detritus	
Carrion	
Labile detritus	

[1] Migratory species move outside of the model domain in parts of the year. All mobile components are able to move within boxes, either due to density-dependent movement, or seasonal migrations, or a mix between these.

[2] The arctic seabirds are parameterized as Brünnich's Guillemot, but represent all seabirds which stay in the northern areas also during wintertime. So the total biomass is much larger than for the guillemots alone.

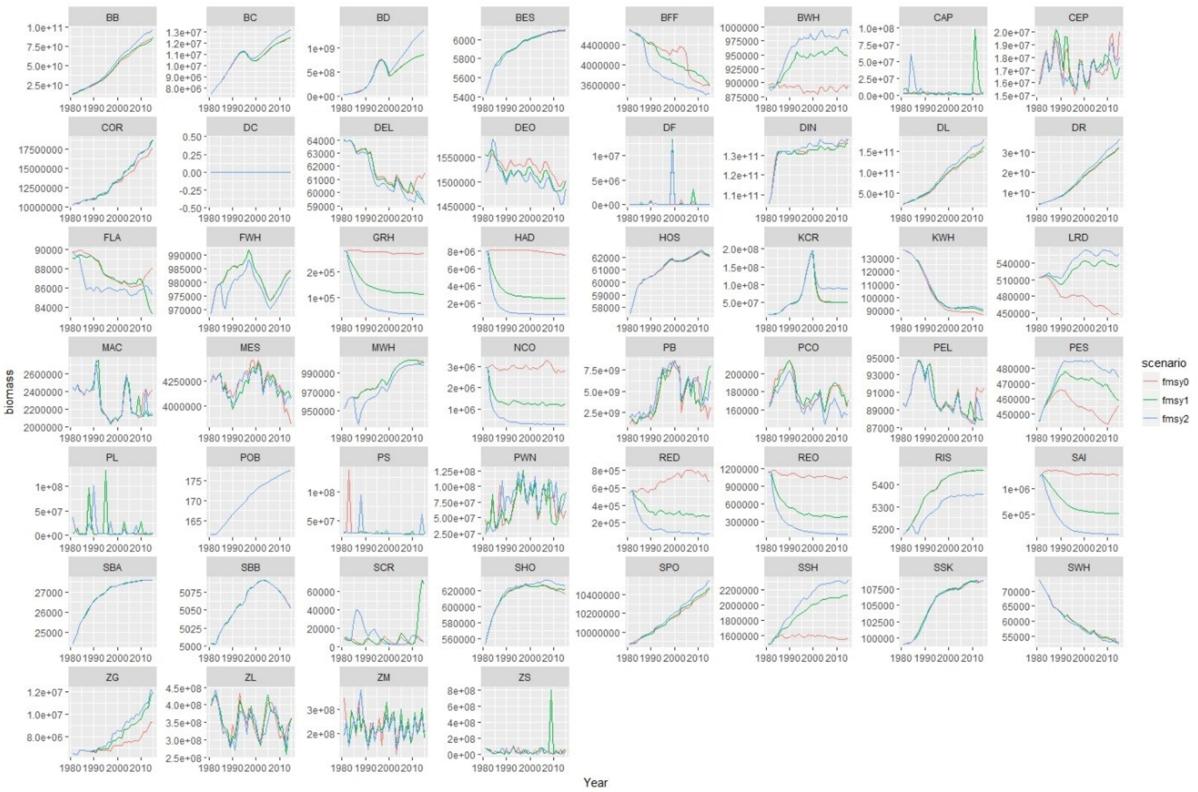

Figure A3-1: Biomass changes for each functional group over three fishing scenarios

12 APPENDIX 4

Table A4-1: Detailed indicator responses at three time points for three fishing scenarios. Figures denote percentage change at each time point compared to the 2015 baseline for each indicator; negative responses indicate biodiversity loss. Dark Red more than 10% reduction, Light Red 0 to -5% reduction, Yellow 0 to 10% gain, Light Green 10% to 50% gain, Dark Green over 50% gain in indicator values.

Scenario		Global Sustainability			Precautionary Fishing			Strict Conservation		
Year		2030	2050	2068	2030	2050	2068	2030	2050	2068
Conservation	LPI	0.38	-6.21	-3.41	4.07	-3.89	-13.30	10.00	-9.30	-10.41
	NNI	-0.59	-0.82	4.45	1.89	-0.57	-1.19	3.75	-3.24	-2.96
	Iconic Abundance	-5.69	-10.64	-14.30	-5.75	-12.43	-16.76	-5.47	-11.63	-15.17
IndiSeas	Total Biomass	7.73	2.82	-4.42	0.71	-0.52	0.50	3.13	1.00	-3.58
	% Pred	4.87	53.89	78.29	34.22	112.72	120.33	85.95	240.82	284.49
	Mean Life Span	2.60	-1.58	-2.43	8.58	3.18	3.84	19.82	20.65	22.11
	TL Community	4.82	6.07	5.67	6.32	15.02	24.10	9.42	25.09	35.90
	Inverse Fishing Pressure	-38.96	2.31	9.08	-15.30	28.00	37.59	NA	NA	NA
	TL Landings	3.04	0.91	-0.09	2.11	0.27	-1.38	NA	NA	NA
Fisheries Ecosystem	PelBioPP	35.76	58.45	31.66	61.14	178.42	229.67	109.81	412.13	545.95
	BioPP	37.45	34.65	-0.73	50.70	115.50	146.19	76.81	230.03	281.24
	DemPel	-13.85	-26.56	-28.18	-7.21	-22.81	-27.94	2.44	-23.35	-28.74
	DemPP	16.88	16.45	-5.75	49.59	114.87	137.37	114.40	293.98	361.00
	PropPel	-8.18	-12.76	-23.12	-4.94	-0.30	-5.85	8.25	24.86	19.71
	Prop Pred	-4.86	1.53	-10.66	10.53	37.01	26.98	47.88	117.61	111.37

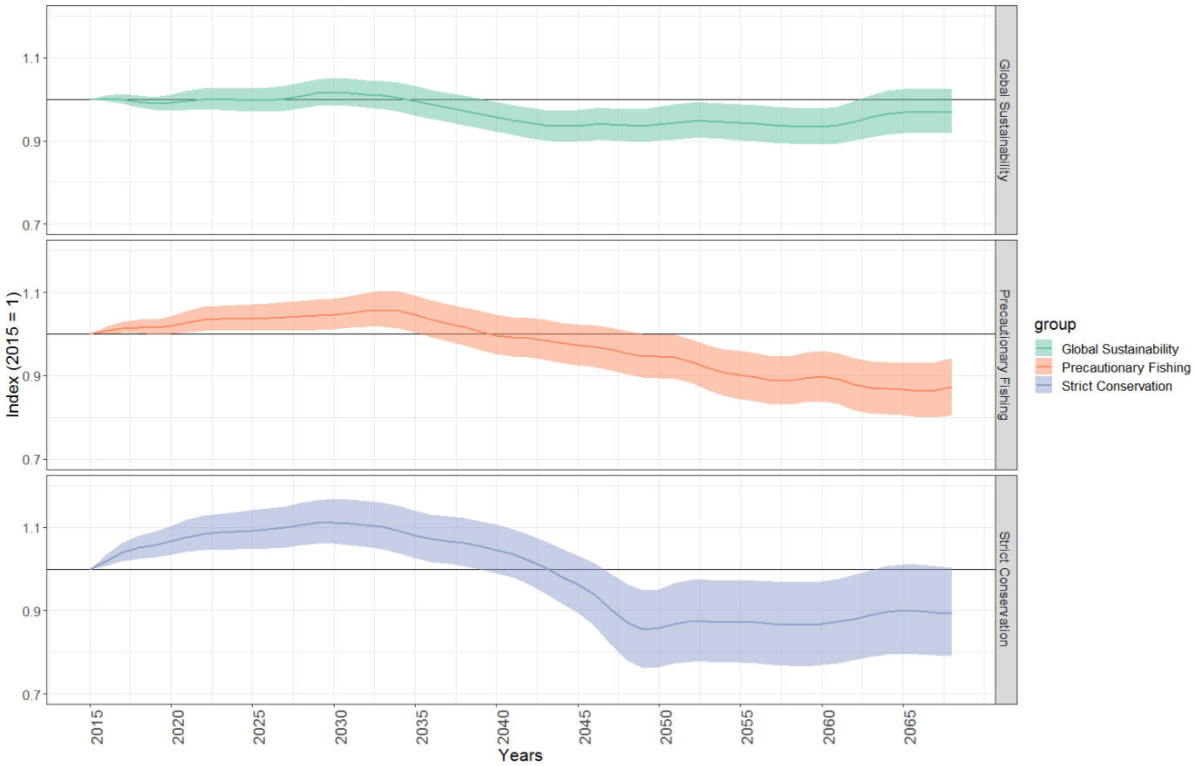

Figure A4-1: Living Planet Index under three fishing scenarios 2015-2068

Figure A4-2: Norway Nature Index under three fishing scenarios 2015-2068

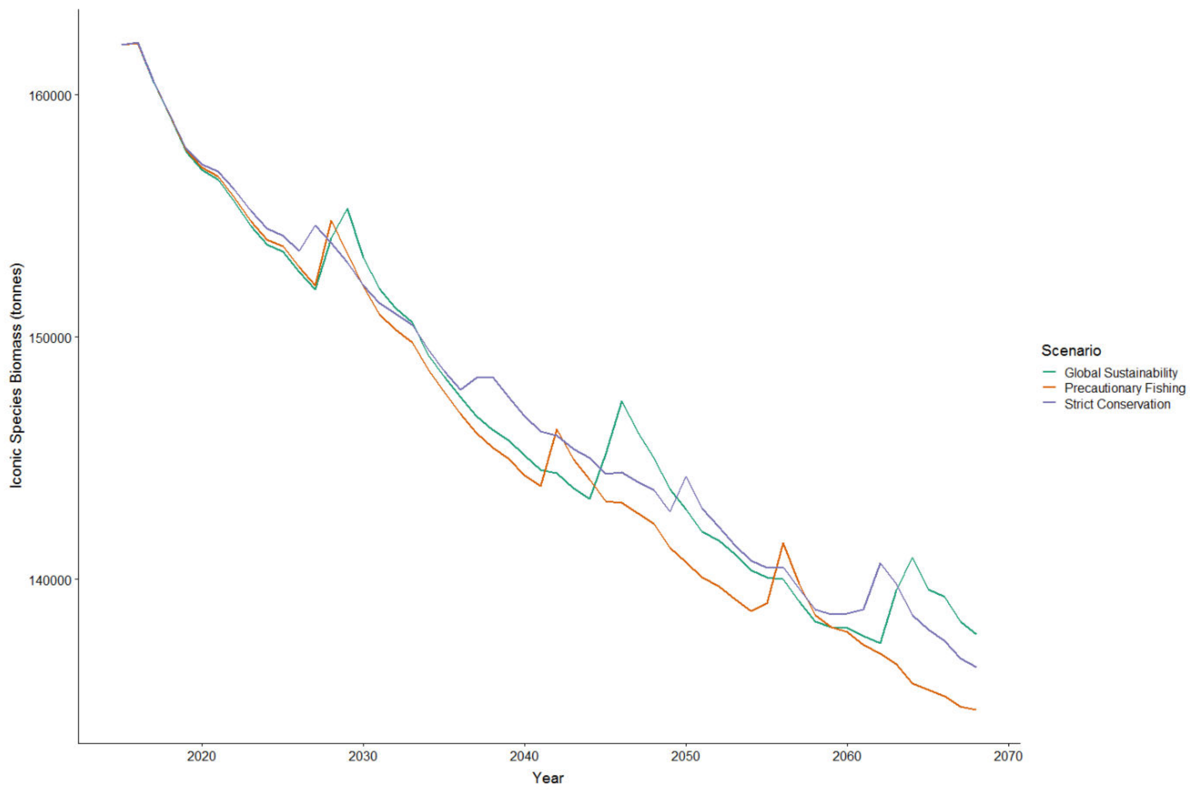

Figure A4-3: Iconic Species Abundance under three fishing scenarios 2015-2068

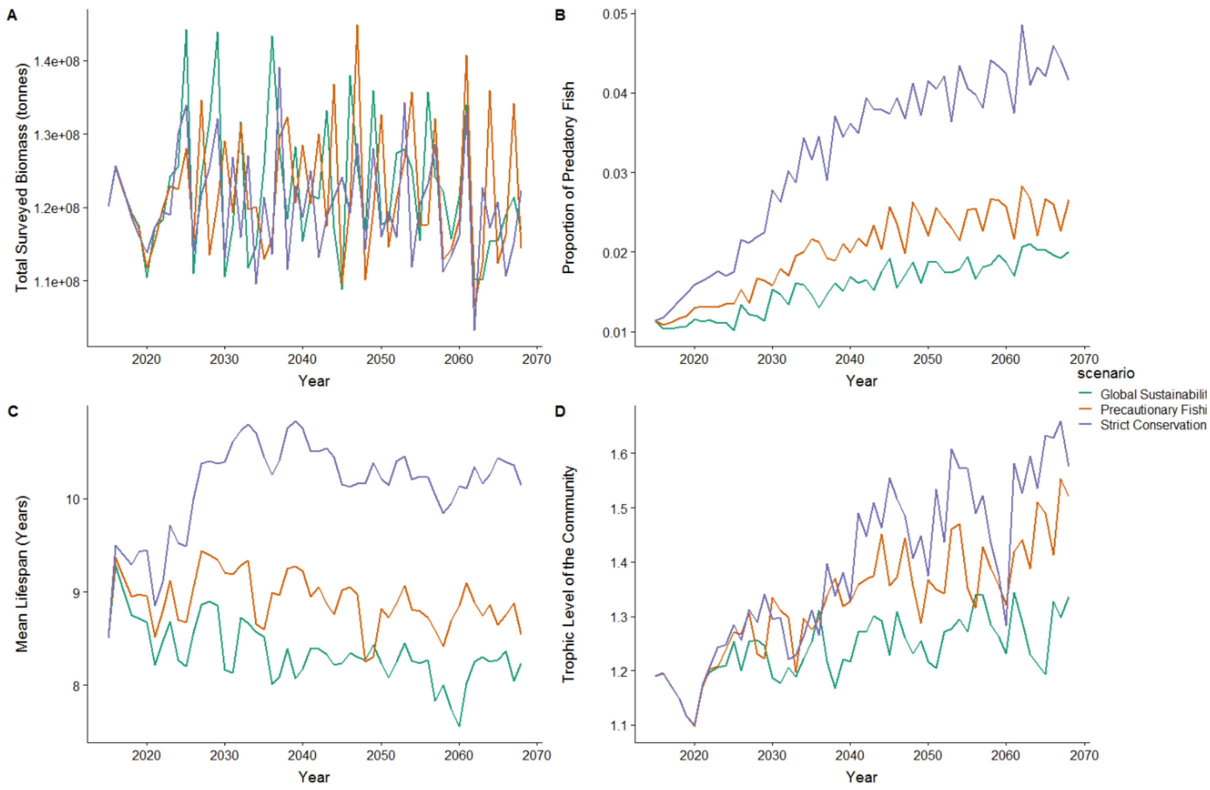

Figure A4-4: IndiSeas biomass based indicators under three fishing scenarios 2015-2068

Figure A4-5: IndiSeas catch based indicators under three fishing scenarios 2015-2068

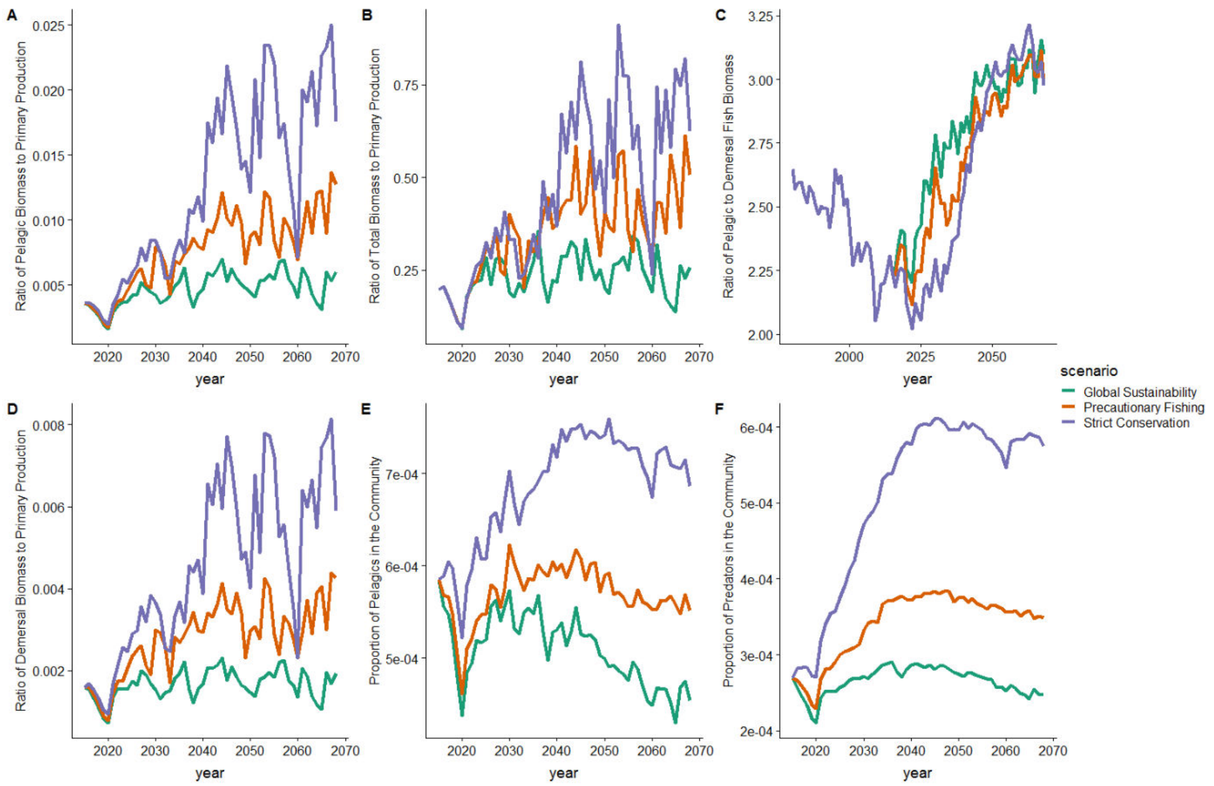

Figure A4-6: Fisheries ecosystem indicators under three fishing scenarios 2015-2068

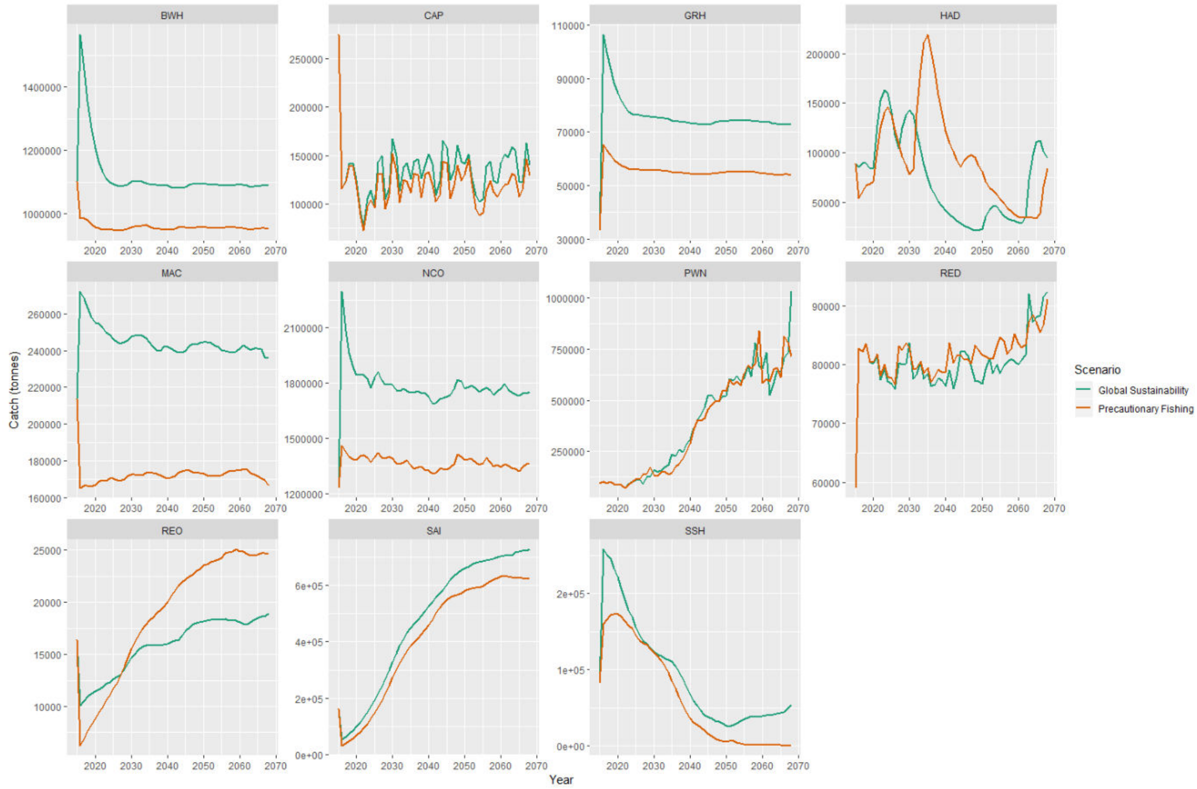

Figure A4-7: Catches for commercial species from 2015-2068 across two fishing scenarios