Estimating unreported catches in Norwegian fisheries

Thomas L Clegg

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2022

UNIVERSITY OF BERGEN

Estimating unreported catches in Norwegian fisheries

Thomas L Clegg

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 27.06.2022

Year: 2022

Title: Estimating unreported catches in Norwegian fisheries

Name: Thomas L Clegg

Print: Skipnes Kommunikasjon / University of Bergen

Scientific environment

The work in this thesis was carried out from 2017 to 2022 in the Fisheries Dynamics research group at the Norwegian Institute of Marine Research (IMR), and through the Department of Biological Sciences at the University of Bergen (UiB), Norway. The Ph.D. was equally funded by IMR and the Norwegian Directorate of Fisheries (FDir).

The research was under the supervision of Dr Kjell Nedreaas (IMR) and co-supervised by Prof Jeppe Kolding (UiB), Dr Geir Blom (FDir), and Dr Kotaro Ono (IMR).

Acknowledgements

I want to start by thanking everyone who has contributed to the Norwegian Reference Fleet programme. The trustful and progressive nature of the programme has made it a pleasure to be a part of and has allowed me to learn and contribute more than I could have expected. I thank all fishers, past and present, and especially the skippers and crews of Hermes and Atlantic who welcomed me on-board to experience data collection first-hand and for the many open discussions about the programme and fisheries. Thank you to Tom Williams for all the support and knowledge you have offered to me throughout the years, and to all the technicians that have helped with any question I had and for maintaining both the relationships with the fishers and the quality of data.

Thank you to my supervisors, Kjell Nedreaas, Jeppe Kolding, Geir Blom, and Kotaro Ono, for the support you have offered throughout my PhD. I wish to thank Kjell, Jeppe, and Geir particularly for giving me the opportunity to study in Norway, and to Kotaro for your extended efforts as a teacher and mentor, all offered at any time of day.

Thank you to everyone in the Fisheries Dynamics group at IMR. You have all helped me along the way in some shape or form. I want to offer a special thank you to Edvin Fuglebakk for everything I’ve learnt from you, from philosophy to sampling design, to Hans Hagen Stockhausen, for support, advice, and friendship, and to Jon Helge Vølstad for your support and guidance during the more difficult times of the PhD.

I thank all my co-authors for their help in improving my scientific writing. I had a look at a few early drafts of manuscripts in recent weeks and realised how much I have improved, all thanks to the sustained efforts to comment and discuss an endless number of versions. This thanks also extends to Ryan Mowat for reading through my thesis and offering some tips and comments.

And finally, thank you Karina. You gave me the space to focus on my PhD, whilst offering unconditional support and the best advice. And you reminded me there’s more to life than fish and stats.

Bergen, April 2022

Abstract

A discard ban for fish was introduced in Norway in 1987, which requires that all commercial catches must be landed and reported. In theory, this regulation creates a full record of total removals from all fisheries. However, exemptions and varying compliance rates create a risk that unreported catches still occur. Estimating unreported catches of all species in multiple fisheries is a large task that is complexified by the many influential factors related to unique fishery regulations, market demands, fishing gear, and species biology. There is therefore a need to standardise the estimation procedure, but this requires compromises that affect the bias and precision variably across individual species which must be understood if results are used as scientific advice. In Norwegian fisheries, the largest source of detailed data on unreported catches comes from the Norwegian Reference Fleet, a group of active fishing vessels that are paid to sample their catches at sea. However, participation in the programme is voluntary, meaning there are uncertainties about how representative the Norwegian Reference Fleet are of the wider fisheries. In such a complex system, it is important to address uncertainties in the entire estimation process, including from sampling data and the estimators used.

The aim of this thesis is to develop standardised estimators for unreported catches in Norwegian fisheries. To identify the current knowledge gaps in Norwegian fisheries, global best practices for estimating unreported catches were collated and applied to Norwegian fisheries. Following from this, two research paths were identified. Firstly, there is a demand to understand the quality of data collected by the Norwegian Reference Fleet. Based on the available data, this was confined to quantifying the representativeness of samples. Secondly, previous studies estimating unreported catches have used bespoke model-based approaches to improve predictive performance, but simple design-based approaches have been applied based on assumptions that have not yet been tested. There is therefore a demand to evaluate the assumptions behind the current design-based estimators.

To evaluate representativeness, the sampling design of the Norwegian Reference Fleet was simulated using reported catches, for which fleet-level information is available. The simulation study identified that nonprobability sampling of vessels in the Norwegian Reference Fleet results in a tendency to overestimate reported catches, but the bias is still within the bounds of expected variation from probability sampling. Representativeness varied greatly across species and years, and there was evidence that the estimators traditionally used for unreported catches may be introducing bias due to assumptions being unmet. These results provide support for the development of improved estimators and consideration of a more conservative estimation of uncertainty. Applying a cluster-based estimator that better describes true variations between sampled vessels produces a more realistic, albeit more uncertain estimate of unreported catches. This is also the case for additional uncertainty incurred from converting numbers of fish to biomass, which must use an additional modelling step due to a lack of information on fish weights. The current methodology for estimating discards in coastal fisheries is restricted by the fishery-level data that is used for extrapolating estimated discard rates. However, current developments in mandatory reporting requirements suggest that future model-based approaches could improve discard estimates. Therefore, an exploratory model was fitted to the sampling data to identify potentially important variables that explain variations in discarding. This model can then inform the variable selection in a future model-based approach when fishery-level data collection is improved.

The estimation methodologies presented in this thesis form the basis of a national routine for estimating unreported catches in Norwegian fisheries. Quantifying the bias of estimators and accounting for additional, important sources of uncertainty provides a standardised design-based estimator for unreported catches in Norwegian fisheries. Predictive performance is now supported by quantitative evidence and further improvements have been identified to optimise estimators in the future such as accounting for rare occurrences and size-based estimates. Furthermore, the lessons learnt throughout this doctoral research highlight the importance of creating a standardised framework for estimating unreported catches. This ensures that improvements are centralised rather than being hidden within individual case studies.

Sammendrag

I Norge ble det innført et utkastforbud for fisk fanget allerede i 1987. I henhold til dette skal all kommersiell fangst føres på land. I teorien oppnår denne forskriften en fullstendig oversikt over totale uttak fra alle fiskerier. Unntak og varierende etterlevelse skaper imidlertid en risiko for at det fortsatt forekommer urapporterte fangster. Å estimere urapporterte fangster av alle arter i flere fiskerier er en stor og komplisert oppgave på grunn av de mange innflytelsesrike faktorene knyttet til unike fiskerireguleringer, markedskrav, fiskeredskaper og artsbiologi. Det er derfor behov for å standardisere estimeringsprosedyren, men dette krever kompromisser som påvirker nøyaktighet og presisjonen i varierende grad på tvers av individuelle arter, og som må forstås hvis resultatene brukes som vitenskapelig råd. I norske fiskerier er Referanseflåten den største kilden til detaljerte data om urapporterte fangster. Referanseflåten er en gruppe aktive fiskefartøyer som får betalt for å ta prøver fra fangstene sine. Siden deltakelse i programmet er frivillig, er det usikkerhet om hvor representativ Referanseflåten er for hele fiskeflåten. I et så komplekst system er det viktig å adressere usikkerhet i hele estimeringsprosessen, inkludert data og estimatorene som brukes.

Målet med denne oppgave er å utvikle standardiserte estimatorer for urapportert fangst i norske fiskerier. For å kartlegge dagens kunnskapshull i norske fiskerier, ble den globale beste praksis for estimering av urapportert fangst sammenstilt og brukt på norske fiskerier. Etter dette ble det definert to forskningsretninger. Det første er nødvendigheten om å forstå kvaliteten på data som samles inn av Referanseflåten. Basert på tilgjengelige data ble dette begrenset til å kvantifisere hvor representativt de innsamlede data er. For det andre har tidligere studier som estimerte urapportert fangst tatt i bruk tilpassede modellbaserte tilnærminger for å forbedre prediktiv ytelse, men noen designbaserte tilnærminger som har blitt brukt er basert på antakelser som ennå ikke er testet. Det er derfor et behov for å evaluere forutsetningene bak designbaserte estimatorer som brukes i dag.

For å vurdere Referanseflåten sin representativitet, ble data innsamlingsdesignet simulert med bruk av rapporterte fangster som er tilgjengelig for hele flåten. Simuleringene viste en tendens til å overestimere rapportert fangst fordi båtene ble ikke valgt ved bruk av sannsynlighet. Likevel er nøyaktigheten fortsatt innenfor rammen av forventet variasjon hvis båtene ble valgt ved bruk av sannsynlighet. Representativiteten varierte sterkt på tvers av arter og år, og det var bevis på at estimatorene som tradisjonelt ble brukt for urapportert fangst, kan innføre unøyaktighet på grunn av at forutsetningene ikke er oppfylt. Disse resultatene gir støtte til utvikling av forbedrede estimatorer og vurdering av en mer konservativ estimering av usikkerhet. Bruk av en klyngebasert estimator som bedre beskriver sanne variasjoner mellom utvalgte fartøyer gir et mer realistisk, om enn mer usikkert estimat av urapporterte fangster. Dette er også tilfellet for ytterligere usikkerhet som følge av konvertering av antall fisk til biomasse, som må bruke et ekstra modelleringstrinn på grunn av mangel på informasjon om fiskevekten. Dagens metodikk for å estimere utkast i kystfiske er begrenset av kvaliteten på dataene på fiskerinivå som brukes for å ekstrapolere estimerte utkastrater. Pågående utvikling i obligatoriske rapporteringskrav tyder imidlertid på at fremtidige modellbaserte tilnærminger kan forbedre estimatene på utkast. Derfor ble en utforskende modell tilpasset prøvetakingsdataene for å identifisere mulige viktige variabler som forklarer grunnene til utkast. Denne modellen kan deretter informere variabelutvalget i en fremtidig modellbasert tilnærming når datainnsamlingen på fiskerinivå forbedres.

Metodene for utkastestimering fremlagt i denne oppgaven kan danne grunnlaget for en nasjonal rutine for å estimere urapportert fangst i norske fiskerier. Å kvantifisere nøyaktigheten til estimatorer og redegjøre for ytterligere viktige kilder til usikkerhet gir en standardisert designbasert estimator for urapporterte fangster i norske fiskerier. Prediktiv ytelse støttes nå av kvantitative bevis og ytterligere forbedringer er identifisert for å optimalisere estimatorer i fremtiden, for eksempel regnskap for sjeldne hendelser og størrelsesbaserte estimater. Erfaringene gjennom denne forskningsoppgave fremhever viktigheten av å skape et standardisert rammeverk for å estimere urapportert fangst. Dette sikrer at forbedringer er sentralisert, i stedet for å være skjult i individuelle casestudier.

List of papers

Paper I
Clegg TL, Kennelly SJ, Blom G, Nedreaas K (2021) Applying global best practices for estimating unreported catches in Norwegian fisheries under a discard ban. Reviews in Fish Biology and Fisheries 31:1–23. https://doi.org/10.1007/s11160-020-09624-w

Paper II

Clegg TL, Fuglebakk E, Ono K, Vølstad JH, Nedreaas K (2022) A simulation approach to assessing bias in a fisheries self-sampling programme. ICES Journal of Marine Science 79:76–87. https://doi.org/10.1093/icesjms/fsab242

Paper III

Clegg, TL, Fuglebakk, E, Ono K. (In prep) Evaluating assumptions behind design-based estimators for unreported catches. Manuscript.

Paper IV

Berg, HSF*, Clegg, TL*, Blom G, Kolding J, Ono K, Nedreaas, K (In press) Discards of cod (*Gadus morhua*) in the Norwegian coastal fisheries: improving past and future estimates. ICES Journal of Marine Science. https://doi.org/10.1093/icesjms/fsac081

** These authors contributed equally to the study.*

Non peer reviewed papers supporting this thesis

Report I
Clegg, T., and Williams, T. 2020. Monitoring bycatches in Norwegian fisheries – Species registered by the Norwegian Reference Fleet. Rapport fra Havforskningen; 2020-8. https://hdl.handle.net/11250/2685855. Accessed 20th April 2022

Report II
Clegg, T. L., Blom, G., Ono, K., and Nedreaas, K. 2021. Estimating the size distribution of reported catches on-board factory vessels – Issues with using data from the production process. ISBN: 9788292075098 https://www.fiskeridir.no/Yrkesfiske/Dokumenter/Rapporter/2021/estimating-the-size-distribution-of-reported-catches-on-board-factory-vessels--issues-with-using-data-from-the-production-process. Accessed 20th April 2022

Contents

1. Introduction

> *'The "best" raising procedure for discards in fishery science is a method everyone talks about but nobody has ever seen it.'*
>
> *(Vigneau 2006)*

1.1 Unreported catches as a global issue

The advent of ecosystem-based fisheries management has changed the way we evaluate the impacts of fishing on the marine environment. Where fisheries management was previously focused on maximising the catches of commercial species, there is now a broader consideration of how fisheries can cause unintended mortality in the wider ecosystem (Pikitch et al. 2004). At the centre of this extended responsibility is the global issue of bycatches (Alverson et al. 1994; Kelleher 2005; Zeller et al. 2017; Pérez Roda et al. 2019; Gilman et al. 2020).

Many commercial fisheries, and particularly those in high-income countries, target species with the highest commercial value. However, no fishing gear is perfectly selective, meaning unwanted bycatches are incurred alongside the commercial catches. The bycatches may be unwanted because of low commercial value or because they are illegal. Either way, these unwanted bycatches are at risk of being discarded at sea, which is widely perceived as a waste of natural resources.

To encourage more selective fishing and the avoidance of unwanted catches, many countries have implemented a discard ban (sometimes referred to as a landing obligation), which requires fishers to land and report everything they catch (Condie et al. 2014; Borges et al. 2016; Karp et al. 2019). This aims to provide a full overview of total removals from the fishery, improving knowledge of the impact that fisheries have on the marine ecosystem. For commercial species, these more complete catch data improve stock assessments by reducing bias in estimates of fishing mortality, stock biomass and biological reference points (Dickey-Collas et al. 2007; Rudd and Branch 2017; Cook 2019; Perretti et al. 2020). For non-commercial species, improved knowledge on catches contributes to a more accurate evaluation of the fishery's broader impact on the ecosystem (Bellido et al. 2011; Collie et al. 2016; Depestele et al. 2019) and is used as evidence in eco-label certification (Giacomarra et al. 2021).

Various regulations and incentives are also available to support a broader discard reduction policy, such as dynamic closures of areas with high bycatch risk (Hazen et al. 2018), adjusting quota allocations (Gullestad et al. 2015), improving fishing gear selectivity (O'Neill et al. 2019), and compensating fishers for landing illegal catches (Gullestad et al. 2015). Whilst these may contribute to an effective discard reduction policy, it is wrong to expect a total elimination of discards and complete catch reporting in selective fisheries (Pitcher et al. 2002; Karp et al. 2019). Firstly, every example of a discard ban worldwide has a list of exemptions related to survivability, commercial importance, or conservation status of species (Condie et al. 2014; Borges et al. 2016; Karp et al. 2019). Secondly, the difficulties of enforcing a discard ban are widely recognised (Kelleher 2005; STECF 2013; Gullestad et al. 2015; James et al. 2019; Karp et al. 2019), given that discarding events occur in a short timeframe over wide expanses of ocean. Alongside this, there is a permanent risk of illegal fishing under any jurisdiction. If these sources of unreported catches (Pitcher et al. 2002; Box 1) are unknown, then it is difficult to evaluate the effectiveness of the discard ban, and the quality of outputs from stock assessments and ecosystem models are compromised.

Discarding is a symptom of selective fishing, so it is valuable to highlight that the issue is of variable concern worldwide. In many Asian and African fisheries, higher utilisation of catches in both subsistence and commercial fisheries means that discarding is not so large of an issue if the fisheries are managed from an ecosystem perspective (Karp et al. 2019). There is nevertheless a bycatch risk for threatened, endangered or protected species (Gray and Kennelly 2018; Karp et al. 2019). Furthermore, if total removals from the fishery are to be known, then unreported catches may still arise from illegal fishing or a lack of reporting requirements.

Selectively removing species from the marine ecosystem, and particularly larger individuals of those species, is recognised to be inconsistent with the principles of ecosystem-based fisheries management (Bellido et al. 2011; Garcia et al. 2012; Borges

Box 1. Sources of unreported catches

Discards are defined as the portion of animals in the total catch which is thrown away or dumped at sea before landing for whatever reason (Kelleher 2005). All catches are at risk of being discarded because of disease, gear damage, or predation. Highly protected species or non-commercial species are sometimes referred to as incidental catches and are at a high risk of being discarded. Fishers may aim to maximize profits by discarding bycatch species with lower value to make more space for catches with higher value when storage is limited, a process known as high-grading (Batsleer et al. 2015). Commercially valuable species are also at risk of high-grading, but in this context the discarded catches are typically smaller individuals which fetch a lower price (Pálsson 2003). Commercial species are also at risk of regulatory discarding if quota is exceeded or if individuals are under the legal minimum size (Batsleer et al. 2015).

Illegal catches are landed but intentionally misreported in contravention of regulations. Whether the catches were taken inside a closed area or season, using illegal fishing methods, or the fisher does not have sufficient quota, illegal catches are intentionally misreported to avoid prosecution. In addition to simply omitting the catches from a report, fish could be concealed as other species or reported to be caught from a different area, resulting in complex combinations of over- and underreporting.

Unmandated catches include catches for which there is no obligation to report explicitly upon landing. A landing obligation may appear unambiguous, but the scope and detail of reporting may result in insufficient knowledge of catches on the level of individual species. As they are of little or no use, they are defined as unreported.

Illegal catches as defined here overlap with illegal, unreported, and unregulated (IUU) fishing (FAO Fisheries Department 2001). However, the ‘unreported’ component of IUU fishing does not consider catches legally unreported (i.e. unmandated catches), or differentiate between discards and illegal catches.

et al. 2016) as it disproportionately impacts food webs and population structures. However, ecosystem-based fisheries management is a broad term with no clear definition or framework (Trochta et al. 2018). There are few examples worldwide of fisheries management that incorporates ecosystem drivers, and stocks are still managed generally on a single-species basis (Skern-Mauritzen et al. 2016). In the context of selective fishing, an ecosystem approach is interpreted as a trade-off between bycatch reduction and utilisation (Kelleher 2005). The importance therefore lies in ensuring good knowledge of unreported catches such that they can be managed sustainably, which requires a quantification for monitoring, whether that be relative or absolute values.

This thesis explores the estimation of unreported catches in the context of Norwegian fisheries. Each fishery has a unique combination of political, economic, and environmental drivers that influence the capture, landing, and reporting of each species. Therefore, estimating unreported catches on a national scale is a large task, but one that can be simplified by developing a standardised methodology that makes the process more manageable whilst allowing for comparisons between fisheries, species, and throughout time. However, a standardised methodology will not perform optimally across all cases. This thesis therefore considers how unique complexities will introduce biases and uncertainties that must be understood to decide how much can be tolerated and identify where improvements can be made.

1.2 Estimating unreported catches

1.2.1 Standardised estimators

A good estimator is one that is simple, easy to implement, performs well on average, and is fit for purpose (ICES 2007a; Cartwright et al. 2016; Stock et al. 2019; Lohr 2021). Standardised estimators for unreported catches commonly use design-based methods. This approach relies on probabilistic rules for selecting samples, such that the theoretical relationship between samples and the population is known and formulae can be defined to estimate values for the population (Lohr 2021). The statistical properties of sampling designs are well understood, with a wide range of methods defined to address almost any population characteristic or practical constraint of sampling (Lohr 2021). These properties of design-based estimators make them a popular choice for routine estimations because their simplicity allows for fast analysis, and comparability across fisheries, species, and years. The simplicity of using defined formulae also helps the transparency of methods, as the assumptions can be stated clearly.

The USA was the first country to develop a standardised routine for estimating bycatches at a national level (Rago et al. 2005; Wigley et al. 2008; NMFS 2011), defining a methodology that covers sampling design, data collection, and estimations procedure. To compare a wide range of fisheries and species, the methodology includes a tier classification system to score the quality of data and available estimation methods (NMFS 2011). The routine also includes a mandatory review exercise every three years (Wigley et al. 2021) to interpret and summarise results from the period and address issues in data collection and estimation, such as selection biases and precision targets. This methodology has been further generalised to be applicable to any jurisdiction worldwide (Kennelly 2020), incorporating the USA's tier classification system for quality control (Kennelly 2020; Benaka et al. 2021).

In addition to standardised routines, specialised best practices have been developed for specific species groups, such as marine mammals (Wade et al. 2021; Moore et al. 2021) and seabirds (Le Bot et al. 2018). When specialising methods, a model-based approach is typically applied to include more species-related information that could improve predictive power or inform advice. In support of specific statistical tools, best practice guidelines have been developed in an ecological or even fisheries context to help improve the statistical quality of estimates (e.g. Martin et al. 2005; Cutler et al. 2007; Bolker et al. 2009; Boulesteix et al. 2012; Harrison et al. 2018; Yan et al. 2021).

1.2.2 Complexity and uncertainty

Unreported catches, like other forms of fisheries advice, are a complex issue with large systemic uncertainties and contrasting values between stakeholders (Johnsen and Eliasen 2011; Dankel et al. 2012). Applying overly complex or non-transparent methods could reduce the ability for non-scientific stakeholders to criticise assumptions or interpret results themselves, impacting the trust in published advice (Cartwright et al. 2016; Sohns et al. 2022). However, ignoring too much complexity can produce the misleading impression that the system is well-understood. Decisions that act on this knowledge are made in false confidence and can lead to damages of trust from stakeholders when intentions are not realised (Cvitanovic et al. 2021).

On the scale of single-species estimates of unreported catches, uncertainties are more readily identifiable and the task of exploring the impact and extent of uncertainties is more manageable (e.g. Breivik et al. 2017). However, if the scope is broadened to a national scale covering multiple fisheries, each with a long list of species, then the approach to handling uncertainty must be modified. To maintain a standardised estimator that is simple enough to be broadly applicable, the complexities introduced by the unique characteristics of different fisheries and species must be ignored to some extent (Rayner 2012). Introducing complexities may only improve estimates in a specific context, such as for a specific species group or only where data are available, but risk reducing the ability to generalise the method.

In the early stages of developing routines for estimating unreported catches, it is therefore useful to identify and address important sources of uncertainty (Janssen et al. 2005; van der Sluijs et al. 2008). Such an approach will help all stakeholders to ‘distinguish the almost certain from the less certain’ (Rosenberg 2007). Quantitative uncertainties can be isolated to evaluate individual sources, but in complex systems, there is also value in qualitative descriptions of uncertainty (Janssen et al. 2005). Whether a lack of data restricts a quantitative evaluation, or the uncertainties are value-laden (e.g. trust; Section 1.3.2), a qualitative description can reduce the risk that uncertainties are disregarded or discredited when not quantified. Ultimately, a good description of uncertainty will build trust early on (Cvitanovic et al. 2021) and help all stakeholders understand the types of uncertainty and where they are found, which will also be of benefit to identify where future improvements may be useful.

1.3 Fisher self-sampling

Catch sampling has been done traditionally by independent scientific observers, which is widely credited as the most reliable data collection method (NMFS 2011; Mangi et al. 2013; Kennelly 2020; Suuronen and Gilman 2020). However, there is no scientific observer programme in Norway due to political reasons (Bowering et al. 2011) and because of the logistical difficulties met when covering coastal fisheries along a complex coastline of long fjords and high mountains ranges. A discard ban is a considerable piece of legislation that fundamentally changes a fishery and questions the role of independent observers in fisheries-dependent sampling. The illegalities surrounding discards and the obligation to land and report all catches mean that where observers have an enforcement role, or even a civic responsibility to report illegal activity, their presence on board will likely impact the fishing behaviour and quality of observations (Cotter and Pilling 2007; Mangi et al. 2013; Ewell et al. 2020). Even where monitoring and scientific roles of observers are separated, concerns about the use of such data may still deter fishers from participating or behaving normally when observed.

In such situations where independent observers may be unsuitable, self-sampling by fishers is emerging as an alternative or supplementary method of fisheries-dependent data collection. There are a broad range of examples where the quality of self-sampling data has been shown to be comparable to more traditional data collection methods (Fox and Starr 1996; Walsh et al. 2005; Starr 2010; Hoare et al. 2011; Lordan et al. 2011; Roman et al. 2011; Kraan et al. 2013; Mion et al. 2015; Bell et al. 2017; Figus and Criddle 2019; Hatlebrekke 2021; Mendo et al. 2022). In fisheries science, the additional benefits of including stakeholders in the scientific process are becoming increasingly recognised (Johnsen and Eliasen 2011; Dankel et al. 2012; Kraan et al. 2013; Mangi et al. 2016, 2018; Calderwood et al. 2021; Mackinson 2022), and is an important aspect of ecosystem-based fisheries management (Garcia and Cochrane 2005; Gullestad et al. 2017; Mackinson and Middleton 2018). Kraan et al. (2013) describes how the benefits go beyond cost-effectiveness and improved sampling coverages to discuss how cooperative research methods can improve the quality and acceptance of scientific advice in contentious settings.

Self-sampling programmes nevertheless possess a unique set of issues concerning data quality due to their nature. Trusting fishers to accurately record data which may negatively impact their livelihood leads to doubts over accuracy and reliability of data (Suuronen and Gilman 2020). However, a quantitative measure of data quality will struggle to capture the dynamics of motivation, quality assurance, and dialogue between fishers and scientists that will influence data quality. Furthermore, the statistical properties of estimates cannot describe the additional benefits gained from cooperative research that improve an understanding and acceptance of scientific advice.

There are some critics of self-sampling that reject the method on the principle that data are inaccurate and unreliable. However, this thesis is positioned on the belief that fishers should be integrated in the scientific process. Recognising the evidence that self-sampling methods have the full potential to provide high-quality data, the criticism should therefore be on the flaws of the specific programme rather than on generalised beliefs. To this end, a quantitative evaluation of data quality is necessary, but it is also useful to address the qualitative aspects of accuracy and reliability that will inform acceptance of data by all stakeholders.

1.3.1 Representativeness

A sample is expected to be representative such that it can be used to describe the population from which it was taken. At any stage of the sampling design where a selection occurs, representativeness is compromised if the selection is biased towards a portion of the population with a certain characteristic, therefore excluding other parts of the population from sampling (Lohr 2021). Selection biases in self-sampling programmes arise from their voluntary[1] nature (Starr 2010; Mangi et al. 2013). Voluntary participation is necessary because it involves fishers spending extra time and effort to sample their catches or fishing activity alongside their normal routines. Representative sampling relies on each unit having a known probability of being selected. If samples can only be taken from vessels that voluntarily participate, then

[1] Voluntary is defined here as participating from one's own choice or consent and does not refer to any compensation or benefits that fishers receive for participating. This distinguishes voluntary self-sampling programmes from mandatory data collection where all fishers are required by law to provide data.

sampling cannot be perfectly probabilistic. Programmes may contact vessels they expect are likely to participate or advertise an open invitation to all vessels. In principle, volunteer sampling cannot be trusted (Lohr 2021), but it is necessary in many scientific fields where probability selection from the population is impossible. To improve biases, elements of judgement sampling can be included, which use a group of experts that review applications and choose vessels they believe to be representative. However, this is only feasible where programmes are willing to restrict the scope of potential applicants.

Best practice guidelines recommend evaluating the representativeness of commercial catch sampling programmes by comparing vessel characteristics of sampled and unsampled vessels in the fishery, alongside characteristics of fishing activity such as geographical spread of activity, average or total fishing effort, and catch composition (ICES 2007a; Cahalan and Faunce 2020). Comparisons are typically done visually, but cluster modelling techniques have also been used to quantitatively evaluate similarities (e.g. Fernandes et al. 2021). However, the nonprobability selection of vessels in self-sampling programmes means the representativeness of data is theoretically unknown. Methods such as judgment selections can provide an assumption of representativeness, but quantitative evidence is necessary to garner acceptance from the scientific community (Kraan et al. 2013).

1.3.2 Reliability

Reliability is formally defined as “the overall consistency of a measure; it has high reliability if it produces similar results when administered under the same conditions” (Lohr 2021). Whilst this formal definition includes selection biases (section 1.3.1), the term is typically used in self-sampling to refer to measurement error and processing error. Measurement error is how much the observation differs from the true value (Lohr 2021). This could be caused by the resolution of the measuring instrument (e.g. fish length measurements being rounded to the nearest centimetre) or its accidental misuse (e.g. weighing scale not tared before measuring). Measurement error can also be intentional, such that the data recorder purposefully records an erroneous observation. Processing error occurs when an accurate observation was made, but for some reason this is not reflected in the dataset (Lohr 2021). For example, a data recorder may mishear the sampler calling out measurements, illegible handwriting may be misunderstood when observations are digitised, or data may be edited during the data cleaning or analysis processes.

The largest concern of measurement error in self-sampling is the risk of bias caused by intentional misreporting of observations (Hoare et al. 2011; Mangi et al. 2013). Because fishers will be directly impacted by the outcome of research that use the data, there is a risk they may manipulate the data to avoid prosecution, protect access to the fishery, or uphold the fishery's public reputation. These factors mean that reliability is a complex issue affected by the broader culture, policy framework, specific fishery regulations, and individual experiences with the scientific process. This risk of intentional measurement error is enough to reject self-sampling data entirely as a method, or at the least expect convincing evidence to support trustworthiness (Suuronen and Gilman 2020). However, Hoare et al. (2011) call for a change in this attitude, as a dogmatic belief of unreliability can no longer be accepted given the success of many self-sampling programmes (Starr 2010; Hoare et al. 2011; Lordan et al. 2011; Kraan et al. 2013) and may actually hamper the progress of self-sampling programmes as a result.

Processing errors in self-sampling are associated with a lack of motivation or time when fishers are instructed to sample alongside regular fishing activity (Mangi et al. 2013), or a lack of resources to aid data collection. These characteristics are therefore related to the broader qualities of the sampling programme, some of which cannot be quantified. The history of trust between fishers and scientists and the motivations for participating will help to evaluate the risk of processing errors. Data quality is also understood by looking at the resources and support available for sampling (e.g. digital data entry), as well as the compensation given to fishers. If these elements of the programme are not clear, then it is difficult for stakeholders or data users to place their trust and evaluate where quality might be of issue.

When addressing reliability, it is vital to distinguish between mandatory self-reporting and voluntary self-sampling. If all fishers have the legal obligation to self-report catches and fishing activity to authorities then compliance will vary dramatically between fishers, particularly if sensitive information such as discards is required (Emery et al. 2019a, b), or if data are used to regulate fishery access (Faunce 2011). In comparison, voluntary self-sampling involves a willingness to participate, meaning compliance is assumed through participation. In fact, accusing willingly participating fishers of fraudulent or poor-quality data based on belief alone can become a self-fulfilling prophecy, as fishers will begin to lose trust and either reduce effort or drop out (Hoare et al. 2011; Ebel et al. 2018). There are nevertheless threats to data reliability in self-sampling programmes, but the risks of these being systemic are much lower than for mandatory self-reporting and are highly dependent on the qualities of the sampling programme, such as motivations, expectations, and dialogue between fishers and scientists.

1.4 Framing unreported catches in a Norwegian context

1.4.1 History and context

The Norwegian Discard Ban (Box 2) is widely seen as a successful implementation of discard reduction policy (Diamond and Beukers-Stewart 2011; Johnsen and Eliasen 2011; Condie et al. 2014; Borges et al. 2016; James et al. 2019; Karp et al. 2019). There is indirect evidence of historical improvements in the status of many Norwegian fish stocks since the implementation of the discard ban in the 1980s (Box 2; Diamond and Beukers-Stewart 2011; Gullestad et al. 2014; Nedreaas et al. 2015). However, rates or levels of discarding are to this day still relatively unknown because there was no direct assessment of the impacts of the discard ban as it was implemented, and routines have not yet been developed for monitoring discards through time. Many stock assessments therefore assume that discarding is negligible in Norwegian fisheries (e.g. ICES 2021) in the absence of evidence.

Whilst incentives can help fishers to improve avoidance tactics or increase the commercial value of bycatches, there remains a permanent risk of capturing species with restricted quota or no commercial value that are unavoidable, which results in the risk of illegal discarding. In addition to this, there is a permanent risk of illegal activity in

Box 2. A brief history of Norwegian fisheries policy and management

(Summarised from Gullestad et al. 2014, 2015, 2017)
A discard reduction policy began in the 1980s when the cod (*Gadus morhua*) population in the Barents Sea was struggling due to historically poor recruitment and high fishing pressure. When a strong year-class was identified in 1983, real-time closures were introduced for areas with high levels of juvenile cod. By 1986, this strong year-class reached the legal minimum landing size, but fishers were still discarding them in favour of older, larger, and therefore more valuable individuals. This practice, known as high-grading, was legal but was nevertheless perceived as a waste of resources and damaging to the future potential of the stock. In 1987, it became illegal to discard dead or dying cod and haddock (*Melanogrammus aeglefinus*).

Over the following decades, governmental white papers have strengthened the importance of reducing unreported catches, and particularly those landed (Nærings- og Fiskeridepartementet 2004; NOU 2019; Fiskeridirektøren 2022). The discard ban has been expanded to include more species, and a wide range of supporting regulations have been introduced to reduce overfishing and unwanted catches through avoidance measures and incentives.

Since the 2000s, Norwegian fisheries management has been transitioning towards an ecosystem approach. In 2009, the new Marine Resources Act was introduced alongside a landing obligation for all species. Alongside a responsibility of maintaining the long-term sustainability of commercially exploited stocks, the act also introduced requirements to conserve biodiversity, which covers all wild, living marine resources. Under the Norwegian approach to ecosystem-based fisheries management, stocks are ranked on their economic importance, which determines their management objectives. Therefore, species of low economic importance must still be monitored to ensure biodiversity and ecosystem function. For each fishery, potential threats are ranked to identify priorities, with a separate table devoted to data-poor species which can sometimes be of conservation concern.

any fishery. The Norwegian Discard Ban has historically been enforced using a mixture of inspections both at sea and on land (Gezelius 2006), but modern control and surveillance technology such as drones, planes, satellites, and on-board cameras are being developed and tested to reduce the need for physical inspections that are limited in coverage (James et al. 2019; Diekert et al. 2021). However, a review of traditional and modern monitoring, control, and enforcement methods by James et al. (2019) concluded that no single method is effective at stopping discarding or illegal fishing, adding that low industry support or poor coverage can reduce effectiveness. Finally, the obligation to land and report all catches means that official catch statistics are of relatively high resolution. There are nevertheless examples where the resolution is too low to evaluate the impact of fishing on individual species. Skates and rays (order: Rajiformes) are notoriously difficult to identify and have been a bycatch of low importance historically (Figueiredo et al. 2020; Amelot et al. 2021), even though many species are threatened by overfishing (Walls and Dulvy 2021). Landings of skates and rays are therefore often grouped together in official catch statistics making it impossible to determine the pressures from fishing on individual species using official catch statistics alone. Similarly, converting unwanted catches into fishmeal or ensilage is an effective utilisation of unwanted catches in Norwegian fisheries. However, there is currently no obligation to report the contents of highly processed products meaning that again there is no information on catches of individual species.

1.4.2 Perspectives

Norwegian fisheries policy (Box 2) aims to reduce all forms or incidental and unwanted mortality of commercially important species, and to protect the biological diversity and ecosystem function with regards to less commercially important species (Gullestad et al. 2017). Therefore, the main use of unreported catch estimates is for input in stock assessments and ecosystem models. However, the aim of reduction also highlights the need to understand the drivers of unreported catches (Northridge et al. 2017) for management to be effective.

The fact that unreported catches contain a mixture of legal and illegal sources is particularly relevant to the context of Norwegian fisheries. Discarding was addressed as a moral issue from the early stages (Box 2; Gullestad et al. 2015). The Norwegian fishing industry has generally high levels of trust in policy decisions (Tirrell 2017) and demonstrate relatively high compliance rates to control and enforcement (Gezelius 2006; Pitcher et al. 2009; Diekert et al. 2021). Such high compliance rates are attributed more to cooperation and voluntary compliance rather than enforcement (Hønneland 2000; Gezelius 2002, 2006; Österblom et al. 2011; Tirrell 2017; Diekert et al. 2021), in addition to a regulatory framework that reduces opportunities to break the law through incentives or reporting requirements (Gezelius 2006). Furthermore, the fishing industry has a long history of participation in the scientific process (Gullestad et al. 2014; **Report I**). Such trust between the fishing industry and institutional stakeholders is built over decades on individual, organisational, and procedural levels, but is vulnerable to being broken in a matter of days (Cvitanovic et al. 2021).

1.4.3 Progress in Norwegian fisheries

The earliest investigations into discarding were centred around gear selectivity. In the 1960s, experimental trawling was used to explore how gear modifications can reduce discarding of cod and haddock in the Barents Sea (Hylen 1965, 1967; Hylen and Smedstad 1974). As the discard ban was implemented, Hylen and Jacobsen (1987) defined a methodology for the estimation of total bycatch of cod in the Barents Sea shrimp fishery, which was later updated to include additional years (Ajiad et al. 2007) and species (Kvamme et al. 2007). This relied on data from control and surveillance monitoring by the Norwegian Directorate of Fisheries and enhanced by data from scientific surveys that overlapped with the fishery.

Early methodologies (see also McBride and Fotland 1996; Dingsør 2001) relied heavily on scientific survey data due to the absence of fisheries-dependent data. However, scientific surveys are almost entirely limited to trawl gears which typically use smaller meshed nets, tow at higher speeds and for shorter times, resulting in different catch compositions. Therefore, catch compositions must be transformed to simulate commercial fishing. This can either be done using theoretical or empirical assumptions.

As computing power increased and modelling methods were developing, discards of cod in the Barents Sea shrimp fishery were estimated using spatial models that made use of control and monitoring data to predict both quantities and risk of bycatches (Aldrin et al. 2011; Breivik et al. 2016, 2017). However, the control and monitoring data used in these case studies (see also Ajiad et al. 2007; Kvamme et al. 2007) are limited to commercial species and the programme is not designed with representativeness, coverage, or sampling rates in mind. Therefore, the ability to standardise such methods across fisheries, species, and years is limited.

The Norwegian Reference Fleet was created in 2000 to gather data on total catches in Norwegian fisheries (Box 3). Data from the programme was first used to estimate bycatches of harbour porpoises (*Phocoena phocoena*) in the coastal gillnet fisheries (Bjørge et al. 2013), where model-based estimators were applied to improve both estimator performance and allow for recommendations of mitigation. Subsequent studies developed the methodology further (Moan 2016; Moan et al. 2020) and expanded the methodology to seals (Moan 2016).

For bycatch estimates of seabirds, methodologies have used a mixture of fisher interviews (Straan et al. 1991; Fangel et al. 2015), the Norwegian Reference Fleet (Fangel et al. 2015; Bærum et al. 2019) and other self-reporting data collection methods (Christensen-Dalsgaard et al. 2019). Model-based estimators have been applied in the most recent studies, which like for marine mammals were used to improve estimator performance (Christensen-Dalsgaard et al. 2019) or to identify dynamics that could inform advice for mitigation (Bærum et al. 2019).

Until this point, all studies focused on single species or species groups. Although, within this scope studies broadened from commercial species to also consider bycatch species of conservation concern. Considering the broader estimation of unreported catches in an entire fishery, earlier studies derived estimates from questionnaire and interview studies with fishers (Hareide and Garnes 2002; Svorken and Hermansen 2014), granting anonymity and asking questions about their perceptions of the broader fishery rather than their own behaviour. The first quantitative study to estimate discards of all species

Box 3. A description of the Norwegian Reference Fleet

The Norwegian Reference Fleet is a group of active fishing vessels that are paid by the Norwegian Institute of Marine Research (IMR) to collect data on their catches and fishing activity. The fleet is divided into a coastal and offshore vessel group, then dividable further by fishing gears for which adapted sampling protocols are defined. Vessels keep a full record of fishing activity and retained catches, with systematic sampling of fishing operations for total catches and individual fish measurements. Coastal vessels record retained and discarded catches explicitly, whilst offshore vessels recorded total catches as one value until 2019, after which catches have been divided explicitly into retained, discarded, and fishmeal. Fishers are also asked to regularly record bycatches of listed rare species. An agreement with control and enforcement authorities ensures that data will not be requested for enforcement purposes, which has not been compromised in the history of the programme.

Four-year contracts are advertised through an open tender process, which list eligibility requirements that aim to ensure representativeness. The vessel selection process aims to select 'typical' vessels in the prioritised fisheries to assume a simple random sample. Applications are assessed by an expert panel, and if multiple vessels are eligible, then the contract is randomly awarded.

Fishers are trained on the electronic data recording system which includes an electronic measuring board and weighing scales linked to a computer tablet display. The software provides sampling protocols for each species, giving warnings if protocols are not followed or potentially erroneous data is entered. Fishers are trained in species identification and are encouraged to verify with IMR taxonomists. Participating vessels are allocated a mentor who keeps in regular contact and visits the vessel at least once a year to check progress and refresh training. Annual general meetings are used a forum for scientists, fishers, and managers to discuss broader aspects such as research, developments, experiences, and concerns.

See **Report I** *for a detailed description of the project aims, fleet structure, and sampling protocols.*

in a fishery was developed in the coastal gillnet fisheries, first for cod (Berg 2019) before being expanded to all species (Berg and Nedreaas 2020).

The more recent studies demonstrate the clearest potential of the Norwegian Reference Fleet for standardising methods of estimating unreported catches. Discards in the coastal gillnet fishery have been quantified using the same design-based estimator for all species, alongside single-species case studies which have developed methods to improve performance. The Norwegian Reference Fleet programme uses standardised sampling protocols (**Report I**), and the importance of monitoring bycatches is listed explicitly in the programme objectives, ensuring longevity of bycatch data collection. However, there is currently limited evidence for the representativeness and reliability of data collected by the Norwegian Reference Fleet, which makes it difficult to evaluate the quality of estimates produced from them.

Case-specific adaptations in previous studies improved estimates for specific species or fisheries, but it means that a standardised methodology has not yet been defined to estimate unreported catches in Norway on a national scale. Therefore, such studies have had to assume bycatch rates based on either expert judgement or by borrowing estimates from similar fisheries (Valdemarsen and Nakken 2002; Nedreaas et al. 2015). Global studies (Alverson et al. 1994; Kelleher 2005; Pérez Roda et al. 2019) have all listed Norway as a specific challenge in assessments due to a lack of available data for individual fisheries. In this situation, they highlight that imputing discard rates for Norway means that results must be interpreted carefully as the uncertainties cannot be captured using the assumptions (Pérez Roda et al. 2019).

1.4.4 Current challenges

Previous assessments of unreported catches in Norwegian fisheries have focused on case studies of species with high commercial value or of high conservation concern and have therefore used bespoke improvements to make better use of the available data or improve estimates by increasing complexity. However, there is still a demand for a standardised approach to estimating unreported catches that is coherent and consistent. There is most potential for such a method in the previous development of design-based methods in the coastal gillnet fisheries using data from the Norwegian Reference Fleet. However, current design-based estimators have been used without a thorough evaluation of the assumptions behind them. If a standardised estimation system is to be deemed accurate and reliable, then these assumptions must be supported by evidence. Furthermore, whilst these methods have potential for estimating unreported catches of species, there is currently no defined approach for size-based estimates of unreported catches, that would indicate high-grading or under-reporting of fish below minimum landing size (Box 1).

In addition to the development of estimators is a demand to understand the quality of data collected by the Norwegian Reference Fleet. The programme was evaluated by an international committee in 2011 to address amongst other points the representativeness and reliability of data (Bowering et al. 2011). The report recommended principally that a comprehensive, analytical review was needed to evaluate the representativeness of individual vessel groups in the fleet, stressing the importance of quantitative evidence. Specific checks of representativeness have been done as preliminary analyses for estimating discards in case study fisheries (e.g. Moan et al. 2020), but the representativeness of Norwegian Reference Fleet data is often assumed (e.g. Bærum et al. 2019; Berg and Nedreaas 2020). The report also recommended secondarily that a quantitative analysis of reliability was needed. However, whilst this is a central theme of this thesis, a quantitative analysis was out of the scope due to the additional complexities of an experimental study design necessary to compare Norwegian Reference Fleet data with another data source of suitable reliability.

2. Aims and objectives

The principle aim of this thesis is to develop standardised estimators for unreported catches in Norwegian fisheries. Standardisation is based on the consistent data collection format by the Norwegian Reference Fleet to develop a methodology that is applicable to all species in the fisheries covered by the sampling programme. To achieve this aim, the following objectives were set:

1. **Identify best practices that can be applied to Norwegian fisheries to help steer the development process.**

In **Paper I** we collate knowledge and experiences from case studies, national initiatives, and international working groups from around the world to define the scope of an estimation study, potential data sources, and appropriate estimation methods. As a result, we produce a list of best practice approaches for estimating unreported catches and an understanding of how they can be applied to the Norwegian system, with consideration of the discard ban. **Report I** further supports this evaluation by providing additional information on the Norwegian Reference Fleet programme in the context of bycatches.

2. **Quantify selection biases in the Norwegian Reference Fleet sampling design to evaluate representativeness.**

The representativeness of total catches is difficult, if not impossible to quantify. An overall evaluation must therefore be built using designing a range of studies that can address specific biases that are likely to affect estimates. **Paper II** contributes to this evaluation by quantifying potential biases arising from the representativeness of the Norwegian Reference Fleet sampling design, specifically the selection of vessels and fishing operations.

3. **Evaluate the assumptions behind the current estimators for unreported catches in Norwegian fisheries**

Paper III evaluates the assumptions behind all available estimators for unreported catches, including those previously applied alongside two proposals. The study uses the Barents Sea longline fishery as a case study to define a methodology for offshore fisheries. We evaluated the assumptions of various candidate estimators by using a resampling method to gain an estimate of the true bias and variance of candidate estimators. This methodology is applied to 30 species in the Barents Sea trawl fishery in Appendix A. In **Paper IV** we adapt this methodology for use in coastal fisheries, where both sampling and population data differ to offshore fisheries. We also address the development of the new mandatory catch reporting system being expanded to coastal vessels. As this new system will improve data availability in coastal fisheries, potentially important variables recorded by the Norwegian Reference Fleet were identified that could be included in a future model-based approach to estimating total discards in the coastal fisheries. Finally, **Report II** presents the results from two pilot studies aimed at developing a size-based estimate of unreported catches by exploring the utility of data collected in the on-board factory production process for describing the size distribution of retained catches.

3. Abstracts of papers

Paper I: Applying global best practices for estimating unreported catches in Norwegian fisheries under a discard ban.

In addition to their role as a fisheries management tool, discard bans can be effective in improving knowledge of total catches via the requirement to land and report all catches. This shifts the focus to understanding the scale of unreported catches in fisheries, rather than only on discards. However, the presence of a discard ban can cause problems with estimation process, as it involves the observation of illegal activities, and the complex sources of unreported catches require a different approach to estimation. The Norwegian discard ban was introduced in 1987 as part of a wider suite of regulatory measures to improve exploitation patterns in commercial fisheries, but a framework for the regular estimation of unreported catches has yet to be established and operationalised. Here, we aim to identify global best practices for estimating unreported catches under a discard ban and assess their applicability to Norwegian fisheries. We approach this in three steps: (1) defining the scope of an estimation, (2) data collection, and (3) the actual procedure for estimation. We discuss how each step can affect the quality of an estimate with regards to accuracy, precision, practical limitations and whether the estimate is fit for purpose. Finally, we provide a list of recommendations for future studies and identify key knowledge gaps and limitations regarding their application to Norwegian fisheries.

Paper II: A simulation approach to assessing bias in a fisheries self-sampling programme.

The hierarchical structure and non-probabilistic sampling in fisher self-sampling programmes makes it difficult to evaluate biases in total catch estimates. While so, it is possible to evaluate bias in the reported component of catches, which can then be used to infer likely bias in total catches. We assessed bias in the reported component of catches for 18 species in the Barents Sea trawl and longline fisheries by simulating 2000 realisations of the Norwegian Reference Fleet sampling programme using the mandatory catch reporting system, then for each realisation we estimated fleet-wide catches using simple design-based estimators and quantified bias. We then inserted variations (e.g. simple random and systematic sampling) at different levels of the sampling design (sampling frame, vessel, and operation) to identify important factors and trends affecting bias in reported catches. We found that whilst current sampling procedures for fishing operations were not biased, non-probabilistic vessel sampling resulted in bias for some species. However, we concluded this was typically within the bounds of expected variation from probabilistic sampling. Our results highlight the risk of applying these simple estimators to all species. We recommend that future estimates of total catches consider alternative estimators and more conservative estimates of uncertainty where necessary.

Paper III: **Evaluating assumptions behind design-based estimators for unreported catches**

Understanding a fishery's impact on the marine ecosystem requires a quantification of total catches, which include unreported catches. For recent years in Norwegian waters, unreported catches have been estimated using data collected by the Norwegian Reference Fleet, a fisher self-sampling programme that regularly gathers data on catches of all species. In this study, we focused on the use of design-based estimators for total catches in offshore fisheries, which have previously been used to estimate discards in the Norwegian coastal gillnet fisheries. After adapting the current methodology to the data available in offshore fisheries, we explored the assumptions behind both unit- and ratio-based estimators, and the effect of ignoring the clustered nature of data. Using a jack-knife resampling method to estimate the true bias in estimates of total catches and associated variability, we found that ignoring the clustered nature of data tended to underestimate the variability, which lead to occurrences where unreported catches were statistically detected when in fact there was too much uncertainty to make such a conclusion. Further validations suggested the assumptions of the cluster unit estimator are likely not met due to a non-random vessel selection that favours more active vessels. We therefore concluded that whilst the cluster ratio estimator performed poorly for the rarest species, this can be seen as an acceptable trade-off to have a reliable estimator for most species.

Paper IV: Discards of cod (*Gadus morhua*) in the Norwegian coastal fisheries: improving past and future estimates

Discarding can be an unknown source of biases and uncertainties in stock assessments. Discarding patterns and quantities vary so a routine methodology for estimating discards is important to give a better picture of total catches, and potentially mortality, in fisheries. Using data from the Norwegian Reference Fleet between 2012 and 2018, this study presents a revised methodology for estimating discards of cod (*Gadus morhua*) in the Norwegian coastal gillnet fisheries which accounts for variations in discarding between vessels and uncertainties in the conversion of numbers to weight discarded. The estimated average discard rate of cod (weight of cod discarded as percentage of total weight caught) is 0.55 % (95 % CI: 0.45-0.70 %), although discard rates in southern areas were an order of magnitude higher than in northern areas. We also present an exploratory analysis of the drivers behind discarding using a random forest regression model. Spatial variations and fishing intensity were identified as the most important drivers of discarding. Results from this study suggest ways in which self-sampled data can be used to estimate discards in Norwegian coastal fisheries, and where accuracy of future estimates can be improved when a higher resolution data collection programme is established.

4. Synthesis and general discussion

Bycatches and discards have emerged over recent decades to become an important issue of fisheries management. As an ecosystem-based approach to fisheries management has developed, the demand for a better understanding of total mortality incurred by fishing activity has led to increased research on estimating unreported catches of all species, rather than only focusing on those with commercial importance. Whilst it is recognised that the Norwegian Discard Ban Package has reduced unwanted catches and discarding in Norwegian fisheries, there is still a need to understand total removals. However, the broad scale of such aims creates issues for delivering high-quality estimates for individual species.

This doctoral thesis presents the development of a standardised method for estimating unreported catches in Norwegian fisheries. The research builds on earlier case studies to develop a methodology which is optimised over all species in a fishery and is extendable to all fisheries covered by the Norwegian Reference Fleet. In developing this methodology, care has been taken to provide quantitative evidence for the decisions made and to communicate uncertainty effectively. **Paper I** contributes to the framing of unreported catches by understanding the issue in the context of Norwegian fisheries, and specifically under the discard ban. After building a picture of the reliability (**Report I**) and representativeness (**Paper II**) of Norwegian Reference Fleet data, case studies were used to develop routine estimation methods (**Paper III**; **Paper IV**; **Report II**) that focus on important sources of uncertainty throughout the estimation process to maintain the credibility of results and transparently communicate the risks across species groups and fisheries.

The ultimate goal for this line of research is the creation of a national monitoring system for unreported catches. This thesis contributes to the goal on a case-specific level of Norwegian fisheries by supporting the choice of estimators for a fishery-scale estimation of unreported catches (**Paper III**; **Paper IV**) and providing evidence for the quality of data on which the estimates stand (**Paper II**). However, there are still aspects that demand further development (**Paper III**; **Report II**) before a standardised monitoring system can be defined. Although the developments of estimation methods presented in this thesis are relevant to Norwegian fisheries, there are still broader themes that are applicable to the wider field of bycatch and discards research. Self-sampling, and particularly reference fleets, are still relatively new methods on an international scale, so understanding the unique and complex issues experienced will be of benefit to other self-sampling programmes in development. On the broadest possible level, the development of standardised estimation systems has now become a goal for many other countries. Therefore, the following discussion evaluates how this thesis contributes to the framework of such a system, identifying where routine elements can be defined and where future developments are necessary.

4.1 The Norwegian Reference Fleet: data quality and acceptance

Evidence for the accuracy and reliability of data on unreported catches from the Norwegian Reference Fleet is mostly provided through descriptions of programme objectives and motivations, sampling protocols, support provided to fishers, and data processing routines (**Report I**). Bjørkan (2011) explains that the scientific community demonstrates their acceptance of data from the Norwegian Reference Fleet programme by authorising its use in the production of scientific knowledge (Williams et al. 2018; ICES 2021), alongside the peer review process of research published in scientific journals. This is demonstrated for unreported catches by peer-reviewed publications estimating discards of highly contentious species, namely seabirds (Fangel et al. 2015; Bærum et al. 2019) and porpoises (Bjørge et al. 2013; Moan et al. 2020). Other stakeholders such as fishers, managers, and policymakers do not often have the power to exercise such authority in the knowledge production process. Scientists do however have the responsibility to provide a transparent communication of the complexities and uncertainties such that results are trusted and accepted by stakeholders (Cartwright et al. 2016; Cvitanovic et al. 2021). However, this type of validation is built on a deeper understanding of the programme by people working closely with the data or interacting with the fishers involved. For those stakeholders that cannot experience this understanding, it is essential that data quality is clearly described, and furthermore that quantitative evidence is provided to address data quality for the wider scientific community (Kraan et al. 2013).

4.1.1 Quantifying representativeness

Paper II uses best practice methods of comparing vessel characteristics and fishing behaviours of sampled and unsampled vessels to build a picture of representativeness. The study offers the most in-depth analysis to date of the representativeness of the Norwegian Reference Fleet. Simulation methods are powerful tools for exploring probability sampling methods that allow for almost any aspect of the sampling design to be tested. The simulations built in **Paper II** focused on how much the historical compositions of the Norwegian Reference Fleet behave like a probability sample, and how this varied across species and through time. In addition to such validations, simulations can be extended as a dynamic tool to determine the minimum sample size needed to achieve the desired precision to reduce sampling costs (Fangel et al. 2015; Li et al. 2019), or identify the most appropriate estimator by testing the assumptions under different scenarios (Diamond 2003; Barlow and Berkson 2012; Cahalan et al. 2015). Such studies use real-world data that rely on high or almost census-level sampling rates to explore the range of scenarios possible. However, simulations can be extended to also generate the underlying fish distributions from which samples are taken (e.g. Kotwicki and Ono 2019). This extension may be particularly useful to address the variations between species identified in **Paper II**.

Previous studies have nevertheless explored various aspects of representativeness. In the context of ling and tusk catches using trawl, gillnets and longline, the Norwegian Reference Fleet was determined only to be representative only of a restricted geographical extent of the fishery (ICES 2007b). A preliminary study in the coastal gillnet fisheries found that representativeness varied depending on the target species and throughout the year (Moan et al. 2020). A study by Fangel et al. (2015) has sometimes been used as evidence for the representativeness of Norwegian Reference Fleet data (Moan et al. 2020; **Paper I**). Seabird bycatches were estimated in the coastal gillnet fisheries by applying the same estimator to two independent sampling designs, producing almost identical estimates. Although the comparison was limited in scope and the analysis was never designed or interpreted as a comparative study, it nevertheless adds to the supporting evidence. Another aim of the Norwegian Reference Fleet programme is to provide data for stock assessments. Hatlebrekke (2021) found that for limited cases, the Norwegian Reference Fleet produced similar trends in catch per unit effort to those generated by fisheries independent research surveys. Whilst this is not directly applicable to representativeness in the context of the wider fishing fleet, this study provides evidence that the Norwegian Reference Fleet sampling design is meeting the intended aims of the project.

The various studies provide evidence that the Norwegian Reference Fleet is representative of the wider fishing fleet in a range of contexts, but there is always a caveat of limitations, such as geographical or temporal (ICES 2007b; Moan 2016; **Paper II**), or in relation to specific species groups (**Paper II**). This is because the sampling design is not defined for the specific purposes of fisheries or species. Instead, the Norwegian Reference Fleet has a broad range of objectives and participating vessels can be active in multiple fisheries. Whilst vessels are selected based on their participation in priority fisheries (**Report I**), contracts are flexible to a vessel altering the harvest strategy or fishing gear over time, and unforeseen circumstances such as vessel refits must be expected over the four-year contract period. Therefore, for any specific case study proposed, the sampling frame and selection probabilities are theoretically undefined, meaning the representativeness of the Norwegian Reference Fleet is theoretically unknown.

Although many studies provide supporting evidence for the representativeness of the Norwegian Reference Fleet in relation to wider fisheries, **Paper II** and **Paper III** both found evidence of a tendency to overestimate total fishery catches using Norwegian Reference Fleet data. The drivers behind this are yet to be determined, as is whether this finding is generalisable to other fisheries. However, Norwegian Reference Fleet vessels are some of the most modern and largest vessels in the fleet and tend to have higher quotas for commercially important species, which may explain higher catch per unit effort. Going beyond speculation to identify the reasons for this overestimation will help to improve future vessel selection methodologies with the aim of mitigating the issue.

Evaluating the representativeness of the Norwegian Reference Fleet is difficult because for nonprobability sampling elements there is no probability theory available to justify the representativeness of samples. It means that for each new case study, the representativeness is theoretically unknown until evaluated. Furthermore, this evaluation is only valid for a period due to the fixed-term nature of contracts and constant developments in the wider fishery. The previous evaluation of the Norwegian Reference Fleet (Bowering et al. 2011) concluded that the fixed-term contracts with vessels create a dynamic risk of representativeness over time, such that further analyses are needed to inform the vessel selection process and ensure representativeness in the future. During the vessel selection process, the same expert judgement panel review all tenders, calling upon additional experts for individual fisheries. Therefore, many aspects of representativeness are standardised across fisheries. However, vessels have a four-year contract for sampling, over which time their representativeness can change. Additionally, vessels are regularly swapped out when contracts expire. This varying representativeness was detected in **Paper II**, demonstrating the constantly changing relationship between the Norwegian Reference Fleet and the fisheries they are supposed to represent, that would be captured automatically in a robust probabilistic sampling design.

4.1.2 Evaluating reliability

The reliability of data collected by the Norwegian Reference Fleet was not quantified in this thesis but should still be addressed. The question of reliability was met throughout this doctoral research process and is one of the most important qualities of self-sampling data (Starr 2010; Hoare et al. 2011; Kraan et al. 2013). **Papers I-IV** and **Report I** (see also Hatlebrekke et al. 2021) address data reliability by offering clear and thorough explanations of the programme structure and sampling protocols of the Norwegian Reference Fleet. Discussion points throughout **Papers I-IV** have identified specific risks to reliability based on expert knowledge, which is helpful for designing future studies that aim to evaluate reliability. On a broader scale, this synthesis is used to reflect on personal experiences gained throughout the research and offer some perspectives that will help to build a fuller picture of data reliability in the absence of quantitative evidence.

The Norwegian Reference Fleet is a collaboration between scientists and fishers to gather high quality data for use in stock assessments and management advice. A questionnaire survey found that fishers participating in the Norwegian Reference Fleet are most motivated most by a '*social responsibility and a wish to strengthen fisheries management*' and '*the opportunity to contribute to marine research*' (Williams et al. 2018). These motivations do not suggest a systemic issue of intentional data manipulation. However, fishers will be less willing to participate in cooperative research projects if scientists doubt the quality and accuracy of data without any relevant evidence (Hoare et al. 2011). Such issues are exemplified in interviews with lobster fishers in Maine, USA, who expressed their frustrations with scientists that do not use the data they collect due to mistrust and do not recognise their knowledge as meaningful or valid (Ebel et al. 2018).

Fishers in the Norwegian Reference Fleet are open about their concerns surrounding data collection. For example, fishers have collectively expressed at the annual general meeting that they are at increased risk of prosecution for bycatches of endangered species during at-sea inspections because of the additional time spent processing catches. Regarding seabirds, there is a common perception from the fishing industry that data will be used to restrict access to the fishery (fishers and IMR staff, personal communications). Allowing fishers to openly express such concerns will build trust in the data which allows issues to be identified and addressed in a constructive manner.

In addition to recognised issues surrounding reliability is the risk of unknown measurement errors that may be impacting data quality. Since the creation of the Norwegian Reference Fleet, 116 vessels have participated (Williams and Gundersen 2021). In 2021, the Norwegian Reference Fleet comprised of 15 offshore vessels and 21 coastal vessels. Moreover, larger offshore vessels may have up to four fishers sampling across trips and shifts. Unsupervised sampling across so many vessels and fishers creates uncertainties about the degree to measurement errors may be occurring. In this respect, quality is maintained through regular visits to vessels. However, even with such quality assurance measures, issues with data quality can still arise for specific cases.

Prior to 2019, catch sampling protocols for offshore trawl vessels required that fishers record the total catch weight of every species, which included the unreported portions (discards and fishmeal). Where possible, the weights should be accurately recorded using the weighing equipment provided. Otherwise, a subsample can be weighed then extrapolated, or the catch weight estimated using expert judgement. During the research process, I discovered that this sampling protocol has been miscommunicated to fishers on trawl vessels. For species processed in the on-board factory, the retained weight was entered instead of the total weight, therefore omitting the unreported portion from the value. This data recording error was identified across many vessels as values were identical to those recorded in their mandatory daily logbooks, and this was also verified by fishers at an annual general meeting. The exact extent of the bias across species and over time is difficult to quantify, so it currently assumed that all observations of total catches for commercial species from trawl vessels are inaccurate. This measurement error results in the inability to estimate unreported catches for any commercial species in trawl fisheries prior to 2019 (Appendix A). Fortunately, the sampling protocols were changed in 2019 to require that fishers record the retained, discarded and fishmeal portions of the catch separately, meaning the issue is no longer applicable.

Another example of risks in data quality assurance regard the bycatch of seabirds. The crew on deck that are processing the catches record bycatches of seabirds in a separate protocol alongside generic catch sampling, and intermittently report these observations to the wheelhouse where they are manually entered into the computer. Poor communication of these additional steps may result in erroneous recording (Tom Williams, personal communication), which would generate unknown measurement errors across vessels. Furthermore, the communication from the deck to wheelhouse creates many opportunities for processing errors to affect the accuracy of data.

Kraan et al. (2013) explain that cooperative research methods in fisheries science are built on values of trust and motivation, neither of which are readily quantifiable, but are better addressed through qualitative methods such as interviews that can capture more fundamental issues in the fishery that may be more long-lived than the quantitative evidence provided for a specific composition of a reference fleet at any given time

(Jacobsen et al. 2012; Ebel et al. 2018; Cvitanovic et al. 2021; Ford and Stewart 2021). Nevertheless, the power of qualitative methods to evaluate data reliability should not neglect the demand for quantitative evidence.

4.1.3 Routine validations of data quality

Making the step from case studies to standardised routines for estimating unreported catches requires that the Norwegian Reference Fleet programme is accepted as a reliable data source in its entirety and not just the small portion investigated. Unlike other national monitoring systems where regional data collection programmes are evaluated in a standardised system (NMFS 2011; Benaka et al. 2021), the Norwegian Reference Fleet spans many fisheries nationally and therefore requires a simplified evaluation because most of the data qualities are consistent across all vessels, fisheries, and species. The transition from case studies to a routine system for the estimation of unreported catches requires a stronger evidence base for the quality of underlying data. Routine estimates should be accompanied by routine assessments of the sampling programme to ensure that quality is consistently upheld (Bowering et al. 2011; Wigley et al. 2021). The USA's Tier Classification System (NMFS 2011) is a good example of how multiple sampling programmes can be compared consistently and transparently, and how quantitative and qualitative descriptions of uncertainty can be combined. The system has been modified for international application (Benaka et al. 2021) but is currently centred around observer sampling programmes. Therefore, small modifications could allow it to include considerations for self-sampling data. The USA standardised bycatch reporting system also includes a three-year review cycle, in which a panel is appointed to run routine analyses (e.g. precision targets) as well as identify potential sources of bias that may be of concern. The simulation model in **Paper II** can be integrated into such a routine assessment by reducing the scope to target the issues of most concern to the representativeness of data, namely the nonprobability selection of vessels.

There is still a demand to validate the data collected by the Norwegian Reference Fleet by comparing data with another source of known reliability. Such a comparison has been done with fisheries-independent data (Hatlebrekke 2021), but it is still important to evaluate whether catch compositions are being reported accurately and reliably for all species, both commercial and non-commercial. Comparing population estimates from two overlapping sampling designs (e.g. Hatlebrekke 2021) will confound all aspects of estimation accuracy, which includes errors due to selection, coverage, measurement, and processing (Lohr 2021). To remove the selection and coverage errors, such a study needs to focus on comparisons of sampling designs using vessels participating in the Norwegian Reference Fleet. For historical studies of reliability, this comparison is limited to cases where Norwegian Reference Fleet vessels have hosted an observer from the Norwegian Directorate of Fisheries or were inspected by the Coast Guard (see **Paper I** for descriptions of these data collection programmes). Such overlaps are rare given the broad coverage of those two programmes, so the statistical power of such a study would be limited.

Bowering et al. (2011) suggested a comparative analysis of samples with and without IMR scientific staff on-board. A historical study using this method with opportunistic comparisons is compromised by low sample sizes. However, an experimental study can be designed with sufficient scope and statistical power to allow for statistical detection of sampling errors and ensure that conclusions are generalisable to the entire programme. When designing such a study, it is helpful to reflect on which type of errors are being identified. For example, intentional misreporting will not occur under supervision. If this is happening for rarer species, the detection of false zeros amongst a high number of true zeros will be a difficult task. Whilst the primary goal is the statistical detection of sampling errors, a descriptive element would also be highly informative. Vigilant supervisors can identify unintentional errors and re-train fishers, creating a constructive improvement of data quality rather than a critical evaluation.

4.2 Estimation procedure

4.2.1 Maintaining simplicity

The design-based estimators previously used for discards in Norwegian coastal fisheries (Fangel et al. 2015; Bærum et al. 2019; Berg and Nedreaas 2020; Moan et al. 2020) are the simplest available and are therefore easy to implement. However, the performance of these simple estimators has been largely unknown, except for comparisons to complex model-based estimators (Breivik et al. 2017; Moan et al. 2020). **Paper III** presents the first quantitative evaluation of performance of various design-based estimators, exploring variations across species and years to identify the estimator which performs best overall. This evaluation is also supported by **Paper II**, in which the simple estimators were found to be biased even when sampling was perfectly probabilistic. **Paper III** and **Paper IV** present various developments of the estimator for unreported catches and discards in the form of accounting for additional sources of uncertainty in estimates. As these developments were all done within the original framework of the original estimator, simplicity is maintained to ensure the proposed estimator is no more difficult to implement, and that methods are readily understandable to a wide range of users and stakeholders. By remaining within the original estimator framework, the new estimator remains fit for the basic purpose of estimating total unreported catches and discards in fisheries, without having to justify any excessive complexities.

4.2.2 Standardising methods

Identifying a consistent and reliable methodology across fisheries is complicated by the diversity in fishery sampling programmes. For example, standardised bycatch reporting in the USA and Australia depends on multiple monitoring programmes operating on federal or regional scales (NMFS 2011; Kennelly 2018). Evaluating the availability and quality of data from these programmes is arguably as large of an exercise as the estimation procedure itself. In this respect, development of routines in Norway benefits from the national sampling coverage of the Norwegian Reference Fleet programme. The sampling coverage of individual fisheries align with national priorities and sampling protocols are kept relatively consistent across vessel groups, with adaptations applied mostly for the gear-specific sampling constraints.

Despite the conveniences of national coverage by the Norwegian Reference Fleet, the unique sampling design results in a distinct set of issues to those of fishery observer programmes. A reference fleet sampling design limits sampling to a small, nonprobability selection of vessels. On the other hand, fishery observer programmes are typically designed with trips as the primary sampling unit (Stratoudakis et al. 1999; Borges et al. 2004; Vigneau 2006), in which multiple fishing operations are observed.

Ignoring the impacts of rejections on bias for the sake of this argument, trip-based design allows more easily for the assumption of simple random sampling of each level, and a sample size that is more controllable in relation to precision or cost goals (Borges et al. 2004; Amande et al. 2012).

Considering the simplicities of a single fisheries sampling programme covering all fisheries, the limited selection of estimators available in the case study of **Paper III** is surprising. In the Barents Sea longline fishery, catch sampling data are abundant and detailed enough to fit the most complex spatiotemporal predictive models available (Yan et al. 2021). However, there is strong evidence that the unit estimators, the simplest cluster-based estimator available, is not suitable due to the poor representativeness of total fishing effort per vessel. This conclusion is also supported by the prior exploration of representativeness in **Paper II**. In the coastal gillnet fishery case study that was used to develop routines for coastal fisheries (**Paper IV**), the limiting factor was the mandatory data collection system used to extrapolate discard rates. This has resulted in the methodology being limited to the unit estimator (Berg and Nedreaas 2020). Fortunately, the use of the cluster unit estimator is justified in this case study because Norwegian Reference Fleet vessels were representative of the wider fishery in terms of both catches of cod and total catches of all species per trip. Therefore, by referring to the development of a standardised method, the estimation process is standardised, but the specific selection of estimator, be it unit-based or ratio-based, just be fully justified by a study on representativeness and assumptions of estimators.

The multispecies case studies in **Paper II** and **Paper III** reveal a clear limitation of the standardised estimator with rarer bycatch species. Such a limitation is expected where a generalised sampling design and single estimator are applied (Lohr 2021), but the issue of rare bycatch estimations is addressed to varying degrees, if at all, across international case studies. For example, the USAs National Bycatch Report (NMFS 2011) describes how precision targets are assessed using aggregate estimates of all species, which informs if sampling rates should be increased. However, such targets are highly case-specific as they depend on how much precision is accepted by managers, and the costs and available funds to increase sampling rates. Alternatively, a list of threatened, endangered or protected species can be defined beforehand and removed from the standardised methodology to avoid a misleading interpretation of results where estimates are most uncertain but stakes are arguably highest (Kennelly 2018).

The decision to provide full results of all species in **Paper III** (and Appendix A) was because of the overall message of the research: that methods are still in development and that improvements are still needed for certain species groups such as rare species. It is therefore that **Paper III** suggests a lower tolerance threshold of both bias and precision where estimates should be improved if management action is justifiable. This is a first step towards developing a structured routine for estimating unreported catches of all species, but the suggested threshold needs to be refined in consideration of the tolerance of end-users and the available resources for improving performance (i.e. further analysis or increased sampling). Furthermore, methods must improve for these rare species for which poor estimation accuracy is not tolerated. Data on rare species are sparse, and are more sensitive to biases from a generalised sampling design, meaning that the zero-generation process may need to be estimated separately or incorporated into the model using zero-inflation methods (Pennington 1996; Martin et al. 2005).

Size-based approaches to estimating unreported catches are not typically incorporated into a routine monitoring system but are nevertheless important for identifying high-grading or the discarding of fish under legal minimum landing size. It is standard practice to use model-based methods for detecting size-based discarding (Allard and Chouinard 1997; Borges et al. 2006), which can then be extrapolated to estimate total discards (Pálsson 2003; Sturludottir et al. 2019). However, despite the availability of methods for size-based estimates of unreported catches, the limitations identified in historical estimates in offshore fisheries are due to a lack of available data. **Report II** summarises two pilot studies that investigated the utility of on-board factory production data for providing information on size-distribution of commercial fish landings. The decision to stop investigating summarised production report data came from the realisation that additional data processing steps necessary to use the data would introduce unacceptable amounts of uncertainty into an estimate. Some of these steps are necessary, such as a conversion from fish weight to length, but others, such as the statistical handling of heavily aggregated data, were deemed excessive. Given that the more accurate estimate of uncertainty in unreported catches is larger than previously understood (**Paper III**; **Paper IV**), the objectives now should be to look for ways in which to reduce uncertainty, not develop methods that will create even more.

4.2.3 Addressing uncertainties

Papers II-IV identified some potentially important sources of uncertainty and explored their impact on estimates of unreported catches. For example, cluster-based sampling by the Norwegian Reference Fleet is recognised as an important source of variance (Aanes and Pennington 2003; Helle and Pennington 2004; Pennington and Helle 2011). However, model-based estimators have only recently began using random effects models as standard practice to account for them (Breivik et al. 2016, 2017; Bærum et al. 2019; Moan et al. 2020), whilst **Paper III** and **Paper IV** are the first case studies that account for cluster-based sampling using a design-based estimator. In the specific field of bycatch estimation, modern developments of estimation methods are almost always done within model-based frameworks. Spatiotemporal models can account for poor sampling coverage, producing more uncertain estimates in areas with low or no sampling (Thorson 2019; Yan et al. 2021). There are even methods available to estimate bycatches when the data are known not to be representative (Authier et al. 2021). With modern advancements in statistical computing, there is now a modelling tool for almost any statistical issue.

The bootstrap procedure was chosen for variance estimation in routines (**Paper III**; **Paper IV**) due to the transparency and flexibility it offers in more complex survey designs. The description of each resampling event in the bootstrap procedure offers transparency to the user to better understand where the uncertainties are found in the system, and the assumptions used to handle them. Moreover, individual components of the bootstrap can be added and modified, as demonstrated in **Paper IV** where the bootstrap was extended to include additional uncertainties in the numbers-to-weight conversion.

Further developments of estimators may reduce uncertainty related to parameter estimations or model structure, but there are still broader uncertainties to consider, some of which cannot be quantified. For example, uncertainties begin when the problem is framed (van der Sluijs et al. 2008). This thesis takes care to frame the issue of unreported catches in the context of Norwegian fisheries. The research focused on the problem of unknown levels of unreported catches in Norwegian fisheries under the assumption that additional mortality is harmful to the marine environment. However, framing the issue as 'removals from the ecosystem' neglects the fact that discarded fish are returned to the ecosystem and provide a food source for scavenging species of fish and seabirds (Depestele et al. 2019; Clark et al. 2020). Handling uncertainties through the involvement of stakeholders such as fishers, managers, and non-governmental organisations (Dankel et al. 2012) is achieved throughout the process of estimating unreported catches, first and foremost through the Norwegian Reference Fleet programme, but also through various channels in the ICES community and Norwegian fisheries management system. Many elements of uncertainty mapping are present in the framework for estimating unreported catches, but the next step is bringing them together in a holistic uncertainty assessment which combines the quantitative and qualitative elements. One such tool is the Numerical Unit Spread Assessment Pedigree (NUSAP) system (Funtowicz and Ravetz 1990; van der Sluijs et al. 2005). This tool extends the classic quantitative descriptions of uncertainty (number, unit, and standard deviation) to include a matrix of uncertainties for each stage of the knowledge production process that allow for a mixture of numerical and descriptive scoring. The Tier Classification System in the USA Standardised Bycatch Report (NMFS 2011) incorporates some of the assessment aspects of the NUSAP system by scoring qualitative judgements about the observer sampling programme.

Uncertainty has a direct impact on the ability to produce actionable results. So far, uncertainty has been discussed in the context of how much to account for. But it is also necessary to consider how uncertainty affects the ability to make actionable conclusions if unreported catches are detected. An estimate of unreported catches (**Paper III**) or discards (**Paper IV**) includes a point estimate and a measure of variability (95 %

confidence interval). If this confidence interval does not include zero, then it is typically stated that an effect has been detected, meaning unreported catches (or discards) have occurred (Nakagawa and Cuthill 2007). However, the de-facto statistical test creates a 5 % probability that the effect is due to chance, such that the conclusion is wrong. When making up to 50 comparisons for all species in a fishery, this risk cannot be ignored. However, this multiple comparisons problem is not widely addressed in multispecies estimations of discards or unreported catches using a design-based approach. The distinction is nevertheless unavoidable as any presentation of variability will encourage the reader to naturally interpret results relative to a reference point, be it reported catches or zero.

On the other hand, there is the risk that excessively accounting for uncertainties results in a paralysis of the decision-making process (Rosenberg 2007; Dankel et al. 2012). However, the uncertainties accounted for in this thesis (vessel-based variation: **Paper III**, **Paper IV**; conversion from numbers to weight: **Paper IV**) cannot be ignored, such that accounting for them creates a realistic picture of uncertainty. However, the ability to act on uncertain knowledge is still possible under the precautionary principle (Sandin 1999; Fischer et al. 2021).

4.2.4 Developing design-based estimators further

The post-stratification system currently used for design-based estimates of unreported catches in all Norwegian studies is based on assumptions that have not yet been tested. The system comprises of a spatial component (statistical area), and a temporal component that includes an annual and seasonal element. There are however small variations of this system, such as in the Barents Sea (**Paper II**; **Paper III**) where the seasonal component comprises of three seasons instead of four (**Paper IV**). The performance of this stratification system was not directly researched in this thesis as it was assumed to be efficient to some degree, but it is still an aspect that can be explored to improve the precision of estimates. Since the development of uncertainty estimation has been prioritised in this thesis, there could be benefits from counteracting some of the added uncertainty by optimising the stratification system. This could be achieved through one of many methodological approaches. An information theoretic approach could propose multiple stratification systems for comparison, an approach typically used in model-based estimators (e.g. Ono et al. 2015; Bærum et al. 2019), but is also achievable for design-based estimators (Stratoudakis et al. 1999). Clustering methods can identify multidimensional classifications of typical fishing activity (e.g. Fernandes et al. 2021). However, for any single study, an objective definition of stratification can improve estimates, but there is no single best approach (Ono et al. 2015). Such methods meet the same obstacles as model-fitting more broadly. Over-fitting of data to optimise stratification may result in high performance for observed fishing activity but may be poor in predicting unobserved activity. Furthermore, complex clustering could improve estimator performance but become too abstract to be interpretable for management advice or application to new case studies.

The potential to enhance an estimate with additional data sources (see **Paper I** for descriptions) should be explored, but developments must be addressed on a case-by-case basis. Any assumptions that the estimate is based on should also be applicable to the new data sources introduced. Efforts spent to understand the quality of data collected by the Norwegian Reference Fleet are obsolete if a new data source is introduced, the quality of which is unknown and therefore assumed. The data source with most potential for enhancing estimates of unreported catches is the Monitoring and Surveillance Service run by the Norwegian Directorate of Fisheries. These data were used to predict historical bycatches in the Barents Sea shrimp fishery (Breivik et al. 2017), assuming representativeness. However, authors emphasised the dangers of generalising the method based on an assumption of representative sampling.

Finally, there was the difficulties met when developing a size-based estimation method in this thesis. The research focused on historical size-based estimates where there were size-based observations of total catches that must be compared with landed catches to infer size-based unreported catches (**Report II**). The report contains a thorough discussion on this topic, so it requires no further contribution here, but it is nevertheless important to list the development need here.

4.2.5 A model-based standardised estimator?

As ICES workshops were developing best practice methods in the 2000s (ICES 2000, 2003, 2007a), the lack of modelling studies was constantly noted. One workshop on discard raising procedures concluded that 'the raising procedure used should be a simple one… However, this does not preclude modelling options that can be available in the future to raise discards…' (ICES 2007a). Since this workshop, there have been various advancements in statistical modelling, many of which have found their way towards applications to fisheries science.

Paper I drew upon best practice methods to conclude that any improvements in estimator performance should '*justify the increase in complexity*' (see also Saltelli 2019; Saltelli et al. 2020). However, considering the process of developing routine methods, this recommendation should really be reversed to state: 'justify the simplicity'. The change may feel semantical and the practice of comparing method remains unchanged, but it shifts the burden of proof to the status quo, rather than each proposed improvement. **Paper III** concludes with a best practice design-based method, but identified some unavoidable issues with design-based methods that are difficult to evaluate in future years. Therefore, it is useful to consider if a model-based approach could improve the standardised estimator.

Models used purely as a predictive tool (in comparison to an exploratory tool; e.g. Bærum et al. 2019; **Paper IV**) are typically associated with complex statistical properties that design-based estimators struggle to capture. For example, nonprobability sampling (Cotter and Pilling 2007), finer-scale spatiotemporal variations (Breivik et al. 2017; Yan et al. 2021), rare events (Martin et al. 2005), or non-linear relationships (Stock et al. 2019, 2020). However, the simplest design-based estimators can be replicated by a linear regression model and still obtain identical point estimates (Lohr 2021). Stratified sampling uses the theoretical framework of a simple linear regression model with categorical variables. The ratio estimator includes an additional slope parameter for a continuous variable (fixing the intercept at zero), and a regression estimator extends to estimate the intercept too. From there, additional variables can be included (multiple linear regression) and the basis is formed for most model-based approaches to estimating discards. These additional variables not only help to explain additional variations in the data and improve predictions, but also help with handling uncertainties when predicting beyond the limits of the sampling data. The clearest example of this is the inclusion of spatiotemporal variation in models (Yan et al. 2021). Instead of devising an *ad hoc*, subjective imputation routine for unsampled strata, spatiotemporal models can interpolate between observations and account for the increasing uncertainty as the distance in space or time increases. This additive process in including variables is exemplified using Norwegian Reference Fleet data by Bjørge et al. (2013) who applied a simple model-based stratified ratio estimator to bycatches of harbour porpoise in the Norwegian coastal gillnet fishery. Moan et al. (2020) developed this model further by including a non-linear relationship between fishing effort and bycatches and considering a hierarchical sampling design.

In any form of standardised estimator, some biases are expected across species or fisheries. It is difficult to explore these biases in a design-based estimator (**Paper III**). In comparison, models are unbiased if the population is adequately described and the model is correctly specified, assuming nothing about how the samples were collected (Cotter and Pilling 2007). Therefore, a well-performed model validation process can evaluate biases and clearly identify model misspecifications. Where assumptions are not met, the model is improvable using a wide range of extensions to address the model misspecifications. Simulation studies (**Paper II**) are nevertheless still a useful tool in addition to model diagnostics to explore the limitations of sampling designs and estimators simultaneously.

Despite the benefits of a well-defined model validation process, such a large amount of control is also a disadvantage for models. Doing a thorough model selection and validation process for each species individually may be an unfeasibly large task, and the process itself may introduce biases from poor model selection methods (Burnham and Anderson 2002). Therefore, a model-based approach would also need some simplifications to make it applicable across all species whilst reducing time and resources to make the process feasible (Appendix B). There is nevertheless a risk that model misspecifications are ignored or overlooked by inexperienced users, and some modelling methods are vulnerable to overfitting. Furthermore, without serious efforts to transparently communicate the decisions, assumptions, and performance evaluations of complex modelling methods, resultant advice risks becoming untrustworthy (Cartwright et al. 2016; Lehuta et al. 2016).

A standardised modelling method will also more readily incorporate a size-based estimation of unreported catches. A size-based estimation of unreported catches has yet to be fully developed due to current limitations in data availability (**Report II**). However, the methodology is expected to follow that developed by Pálsson (2003), with uncertainty estimated using a methodology subsequently proposed by Sturludottir et al. (2019). An additional benefit of this approach is the fact that the size-based estimate is produced independent of the standardised estimator, creating the opportunity to compare them as a form of validation.

4.3 General perspectives on estimating unreported catches

Data collection by the Norwegian Reference Fleet is going through a constant process of development and refinement. Of most relevance to this thesis is the new sampling protocols for offshore vessels since 2019 which now record retained catches, discards, and fishmeal explicitly. Furthermore, incidental catches of corals and sponges are now recorded across vessels, which contributes towards sustainability certification from the Marine Stewardship Council. Vessels are also encouraged to record all incidental catches of the rarest species[2], not just for sampled fishing operations. These extra reports are not yet reliable enough to be included in analyses because efforts are not consistent across all vessels, but the information is nevertheless useful for the early detection of trends and can be improved in the future. The cooperative nature of data collection means that expansions or improvements of sampling protocols cannot be rushed, as trust and understanding take time to build (Cvitanovic et al. 2021), and workload needs to be

[2] There is no defined list of rare species that should be recorded. Rather, relevance is open to interpretation and many fishers offer these observations during discussions with IMR staff.

tested slowly such that fishers do not find themselves overloaded and at risk of compromising data quality.

The historical limitation in offshore Norwegian fisheries of estimating unreported catches, and not discards specifically, is a relatively unique situation which has its advantages and disadvantages. Reporting total catches grants a degree of protection to individual fishers by not explicitly recording the unreported portions. The resultant method of inferring unreported catches by comparing estimated total catches with reported catches adds a layer of uncertainty and obscures the incriminatory data generated by individual vessels which will encourage more accurate reporting. However, an estimate of total unreported catches in a fishery offers no information on either the sources of unreported catches, or an understanding of the drivers behind them. The transition by offshore vessels in the Norwegian Reference Fleet to recording discards explicitly will address some of this unknown, but it then neglects the other sources of unreported catches, particularly illegally landed catches which are recognised as a major source of unreported catches in some Norwegian fisheries (Blom et al. 2020). Therefore, estimating unreported catches is an effective way of screening fisheries to identify high-risk fisheries that would prompt further analysis or data collection. If estimates of unreported catches are only needed to adjust reported catches used in stock assessments, then the sources or drivers are irrelevant. However, if the aim is also to reduce unreported catches, then sources and drivers (e.g. Breivik et al. 2016; Northridge et al. 2017; Bærum et al. 2019; **Paper IV**) should be understood to inform decisions.

Beyond a firm understanding of sampling design, a broader familiarity with a sampling programme is useful to ensure that data are used correctly. For example, it is important to know if sampling protocols changed over time to avoid misinterpreting trends. In reference fleet programmes, it is particularly important to understand variations in data across vessels. As vessels are not sampled probabilistically and remain in the reference fleet over time, data users should be aware of variations in fishing activity that may affect representativeness. In the specific context of the Norwegian Reference Fleet, keeping in close contact with IMR staff that are responsible for individual vessels will help identify such trends that may not be obvious in the data. Finally, as this doctoral research has demonstrated, an intimate knowledge of the data will help to identify data issues (e.g. Appendix A) that would risk passing undiscovered if routines were applied without offering much attention to the underlying data.

Mandatory reporting in Norwegian fisheries is also going through a process of heavy development which will change our knowledge of unreported catches. The expansion of electronic logbooks to all Norwegian vessels has the potential to improve estimation methods (**Paper IV**), whilst the movement towards automatic data collection at the closest possible stage after capture will redefine the entire issue of unreported catches (NOU 2019; **Report II**). Under the current reporting system, there are many opportunities to misreport catches (Box 1), and therefore many methods for estimating them. Moving the reporting system to the earliest possible stage focuses the enforcement and monitoring and creates for a single point of verification to evaluate whether reported catches are accurate and reliable. The system removes unmandated catches entirely as the entire catch is automatically recorded, whilst both discards and illegal catches are then identifiable through comparisons with landings and sales reports. This development is in a positive direction and will reduce much of the complexity that is currently met when estimating unreported catches. However, care must be taken to ensure continuity of knowledge throughout the transition phase and new methods will need to be defined once fully automatic monitoring of catches is operational (**Report II**).

4.4 Conclusion

A routine system for estimating unreported catches in Norwegian fisheries is only just beginning to take shape. Generalised methods that were only recently defined (Berg and Nedreaas 2020; **Paper III**) have now been evaluated to offer some insight into their performance (**Paper III**; **Paper IV**). The long time series of total catch data collected by the Norwegian Reference Fleet is also in the early stages of being explored and understood. As more studies use the data, the understanding of data quality improves on a general, fishery, and species level. Defining reliable routines is complicated by the cooperative research with fishers which introduces additional elements of uncertainty, many of which are qualitative and not readily accepted. However, transparency is vital here to be able to identify where strengths and weaknesses are in the sampling programme to avoid generalised discrediting of data.

The process of developing the wider structure around a routine estimation system in Norway can benefit from the long history of the USA's standardised bycatch reporting system. The USA's path began in 1996 with the introduction of a legal mandate to monitor bycatches, but it took 15 years before the first official report was published (NMFS 2011), and development is still a continual process (Wigley et al. 2021). The decades of effort to create and maintain such a large-scale system reflects the complexities that need to be addressed, but also paves the way for other nations to learn and adapt when developing a similar system. For example, Tier Classification System used to evaluate data quality in USA fisheries was developed for the National Bycatch Report (NMFS 2011), but has since been applied to Australian fisheries (Benaka et al. 2021).

A piecemeal approach to estimating unreported catches upon request could be done today using the methods presented in this thesis. However, I recommend against such an approach because there is still much scope for developing both data collection and estimation methods. Addressing these in the context of many routine estimation studies occurring simultaneously would scatter the developments across many sources and lead to bespoke solutions that may help to improve the specific case, but reduce the comparability and continuity across fisheries, species, and years. These issues are further exacerbated by the fact that routine estimates are often published as working documents or reports, making any developments less searchable for others to implement. Therefore, I advocate a more strategic and centralised approach reflecting the process in USA fisheries. Addressing developments in devoted studies would help to maintain standardisation and more easily communicate improvements to the researchers applying methods routinely.

In the consideration of data quality, there is still scope for improving the acceptance of self-sampling data from the Norwegian Reference Fleet. This thesis addresses the recommendations from Bowering et al. (2011) by quantitatively evaluating representativeness and providing suggestions on how these can be generalised. However, based on my experience with the Norwegian Reference Fleet programme, the greatest risk to data acceptance is the trust from the wider scientific community that fishers are not intentionally misreporting their observations. This is a common concern across self-sampling programmes but should a priority for the Norwegian Reference Fleet given the lack of quantitative evidence.

References

Aanes S, Pennington M (2003) On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. ICES Journal of Marine Science 60:297–303. https://doi.org/10.1016/S1054-3139(03)00008-0

Ajiad A, Aglen A, Nedreaas K, Kvamme C (2007) Cod bycatches in the Barents Sea shrimp fishery during 1983-2005. NAFO/ICES Pandulus assessment group meeting NAFO SC Doc. 07/86

Aldrin M, Mortensen B, Storvik G, et al (2011) Improving management decisions by predicting fish bycatch in the Barents Sea shrimp fishery. ICES Journal of Marine Science 69:64–74. https://doi.org/10.1093/icesjms/fsr172

Allard J, Chouinard GA (1997) A strategy to detect fish discarding by combining onboard and onshore sampling. Canadian Journal of Fisheries and Aquatic Sciences 54:2955–2963. https://doi.org/10.1139/f97-180

Alverson DL, Freeberg MH, Murawski SA, Pope JG (1994) A global assessment of fisheries bycatch and discards. FAO Technical paper 339. Rome, Italy. https://www.fao.org/3/T4890E/T4890E00.htm. Accessed 20th April 2022

Amande MJ, Chassot E, Chavance P, et al (2012) Precision in bycatch estimates: the case of tuna purse-seine fisheries in the Indian Ocean. ICES Journal of Marine Science 69:1501–1510. https://doi.org/10.1093/icesjms/fss106

Amelot M, Batsleer J, Foucher E, et al (2021) Evidence of difference in landings and discards patterns in the English Channel and North Sea Rajidae complex fishery. Fisheries Research 242:106028. https://doi.org/10.1016/j.fishres.2021.106028

Authier M, Rouby E, Macleod K (2021) Estimating cetacean bycatch from non-representative samples (I): a simulation study with regularized multilevel regression and post-stratification. Frontiers in Marine Science 8:1459. https://doi.org/10.3389/fmars.2021.719956

Bærum KM, Anker-Nilssen T, Christensen-Dalsgaard S, et al (2019) Spatial and temporal variations in seabird bycatch: Incidental bycatch in the Norwegian coastal gillnet-fishery. PLoS One 14:e0212786. https://doi.org/10.1371/journal.pone.0212786

Barlow PF, Berkson J (2012) Evaluating methods for estimating rare events with zero-heavy data: a simulation model estimating sea turtle bycatch in the pelagic longline fishery. Fishery Bulletin 110:344–360

Batsleer J, Hamon KG, van Overzee HMJ, et al (2015) High-grading and over-quota discarding in mixed fisheries. Reviews in Fish Biology and Fisheries 25:715–736. https://doi.org/10.1007/s11160-015-9403-0

Bell RJ, Gervelis B, Chamberlain G, Hoey J (2017) Discard estimates from self-reported catch data: an example from the U.S. northeast shelf. North American Journal of Fisheries Management 37:1130–1144. https://doi.org/10.1080/02755947.2017.1350219

Bellido JM, Santos MB, Pennino MG, et al (2011) Fishery discards and bycatch: solutions for an ecosystem approach to fisheries management? Hydrobiologia 670:317–333. https://doi.org/10.1007/s10750-011-0721-5

Benaka LR, Chan AN, Kennelly SJ, Olsen NA (2021) Using a tier classification system to evaluate the quality of bycatch estimates from fisheries. Reviews in Fish Biology and Fisheries 31:737–752. https://doi.org/10.1007/s11160-021-09670-y

Berg HSF (2019) Estimation of discard of cod (*Gadus morhua*) in Norwegian gillnet fisheries. Masters thesis, University of Bergen. https://hdl.handle.net/1956/20310. Accessed 20th April 2022

Berg HSF, Nedreaas K (2020) Estimering av utkast i Norsk kystfiske med garn 2012-2018. Fisken og havet 2021-1. Institute of Marine Research, Brage IMR https://hdl.handle.net/11250/2725144. Accessed 20th April 2022

Bjørge A, Skern-Mauritzen M, Rossman MC (2013) Estimated bycatch of harbour porpoise (*Phocoena phocoena*) in two coastal gillnet fisheries in Norway, 2006–2008. Mitigation and implications for conservation. Biological Conservation 161:164–173. https://doi.org/10.1016/j.biocon.2013.03.009

Bjørkan M (2011) Fishing for Advice: the case of the Norwegian Reference Fleet. Doctoral thesis, University of Tromsø. UiT Munin https://hdl.handle.net/10037/3770. Accessed 20th April 2022

Blom G, Misund R, Aarbakke B, et al (2020) Artssammensetning og lengdefordeling i fisket med småmasket trål etter målartene øyepål og kolmule i perioden 2014-2019. Fiskeridirektoratet. https://www.fiskeridir.no/Yrkesfiske/Dokumenter/Rapporter/2020/artssammensetning-og-lengdefordeling-i-fisket-med-smaamasket-traal-etter-maalartene-oeyepaal-og-kolmule-i-perioden-2014-2019. Accessed 20th April 2022

Bolker BM, Brooks ME, Clark CJ, et al (2009) Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology and Evolution 24:127–135. https://doi.org/10.1016/j.tree.2008.10.008

Borges L, Cocas L, Nielsen KN (2016) Discard ban and balanced harvest: a contradiction? ICES Journal of Marine Science 73:1632–1639. https://doi.org/10.1093/icesjms/fsw065

Borges L, Zuur AF, Rogan E, Officer R (2004) Optimum sampling levels in discard sampling programs. Canadian Journal of Fisheries and Aquatic Sciences 61:1918–1928. https://doi.org/10.1139/f04-138

Borges L, Zuur AF, Rogan E, Officer R (2006) Modelling discard ogives from Irish demersal fisheries. ICES Journal of Marine Science 63:1086–1095. https://doi.org/10.1016/j.icesjms.2006.03.017

Boulesteix A, Janitza S, Kruppa J (2012) Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. WIREs Data Mining Knowledge and Discovery 2:493–507. https://doi.org/10.1002/widm.1072

Bowering R, Storr-Paulsen M, Tingley G, et al (2011) Evaluation of the Norwegian Reference Fleet. Institute of Marine Research, Bergen. Brage IMR. http://hdl.handle.net/11250/116704. Accessed 20th April 2022

Breivik ON, Storvik G, Nedreaas K (2017) Latent Gaussian models to predict historical bycatch in commercial fishery. Fisheries Research 185:62–72. https://doi.org/10.1016/j.fishres.2016.09.033

Breivik ON, Storvik G, Nedreaas K (2016) Latent Gaussian models to decide on spatial closures for bycatch management in the Barents Sea shrimp fishery. Canadian Journal of Fisheries and Aquatic Sciences 73:1271–1280. https://doi.org/10.1139/cjfas-2015-0322

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag New York. https://doi.org/10.1007/b97636

Cahalan J, Faunce C (2020) Development and implementation of a fully randomized sampling design for a fishery monitoring program. Fishery Bulletin 118:87–99. https://doi.org/10.7755/FB.118.1.8

Cahalan JA, Gasper J, Mondragon J, Trenkel V (2015) Catch estimation in the federal trawl fisheries off Alaska: a simulation approach to compare the statistical properties of three trip-specific catch estimators. Canadian Journal of Fisheries and Aquatic Sciences 72:1024–1036. https://doi.org/10.1139/cjfas-2014-0347

Calderwood J, Marshall CT, Haflinger K, et al (2021) An evaluation of information sharing schemes to identify what motivates fishers to share catch information. ICES Journal of Marine Science fsab252. https://doi.org/10.1093/icesjms/fsab252

Cartwright SJ, Bowgen KM, Collop C, et al (2016) Communicating complex ecological models to non-scientist end users. Ecological Modelling 338:51–59. https://doi.org/10.1016/j.ecolmodel.2016.07.012

Christensen-Dalsgaard S, Anker-Nilssen T, Crawford R, et al (2019) What's the catch with lumpsuckers? A North Atlantic study of seabird bycatch in lumpsucker gillnet fisheries. Biological Conservation 240: 108278. https://doi.org/10.1016/j.biocon.2019.108278

Clark BL, Vigfúsdóttir F, Jessopp MJ, et al (2020) Gannets are not attracted to fishing vessels in Iceland—potential influence of a discard ban and food availability. ICES Journal of Marine Science 77:692–700. https://doi.org/10.1093/icesjms/fsz233

Collie JS, Botsford LW, Hastings A, et al (2016) Ecosystem models for fisheries management: Finding the sweet spot. Fish and Fisheries 17:101–125. https://doi.org/10.1111/faf.12093

Condie HM, Grant A, Catchpole TL (2014) Incentivising selective fishing under a policy to ban discards; lessons from European and global fisheries. Marine Policy 45:287–292. https://doi.org/10.1016/j.marpol.2013.09.001

Cook RM (2019) Inclusion of discards in stock assessment models. Fish and Fisheries 20:1232–1245. https://doi.org/10.1111/faf.12408

Cotter AJR, Pilling GM (2007) Landings, logbooks and observer surveys: improving the protocols for sampling commercial fisheries. Fish and Fisheries 8:123–152. https://doi.org/10.1111/j.1467-2679.2007.00241.x

Cutler DR, Edwards TC, Beard KH, et al (2007) Random forests for classification in ecology. Ecology 88:2783–2792. https://doi.org/10.1890/07-0539.1

Cvitanovic C, Shellock RJ, Mackay M, et al (2021) Strategies for building and managing ‘trust’ to enable knowledge exchange at the interface of environmental science and policy. Environmental Science & Policy 123:179–189. https://doi.org/10.1016/j.envsci.2021.05.020

Dankel DJ, Aps R, Padda G, et al (2012) Advice under uncertainty in the marine system. ICES Journal of Marine Science 69:3–7. https://doi.org/10.1093/icesjms/fsr179

Depestele J, Feekings J, Reid DG, et al (2019) The Impact of Fisheries Discards on Scavengers in the Sea. In: Uhlmann SS, Ulrich C, Kennelly SJ (eds) The European Landing Obligation: Reducing Discards in Complex, Multi-Species and Multi-Jurisdictional Fisheries. Springer, Cham, pp 129–162. https://doi.org/10.1007/978-3-030-03308-8

Diamond B, Beukers-Stewart BD (2011) Fisheries discards in the North Sea: waste of resources or a necessary evil? Reviews in Fisheries Science 19:231–245. https://doi.org/10.1080/10641262.2011.585432

Diamond SL (2003) Estimation of bycatch in shrimp trawl fisheries: a comparison of estimation methods using field data and simulated data. Fishery Bulletin 101:484–500 https://spo.nmfs.noaa.gov/content/estimation-bycatch-shrimp-trawl-fisheries-comparison-estimation-methods-using-field-data. Accessed 20th April 2022

Dickey-Collas M, Pastoors MA, van Keeken OA (2007) Precisely wrong or vaguely right: simulations of noisy discard data and trends in fishing effort being included in the stock assessment of North Sea plaice. ICES Journal of Marine Science 64:1641–1649. https://doi.org/10.1093/icesjms/fsm155

Diekert F, Nøstbakken L, Richter A (2021) Control activities and compliance behavior—Survey evidence from Norway. Marine Policy 125:104381. https://doi.org/10.1016/j.marpol.2020.104381

Dingsør GE (2001) Estimation of discards in the commercial trawl fishery for cod. Masters thesis, University of Bergen. Bergen Open Access Research Archive, http://hdl.handle.net/1956/1767. Accessed 20th April 2022

Ebel SA, Beitl CM, Runnebaum J, et al (2018) The power of participation: Challenges and opportunities for facilitating trust in cooperative fisheries research in the Maine lobster fishery. Marine Policy 90:47–54. https://doi.org/10.1016/j.marpol.2018.01.007

Emery TJ, Noriega R, Williams AJ, Larcombe J (2019a) Changes in logbook reporting by commercial fishers following the implementation of electronic monitoring in Australian Commonwealth fisheries. Marine Policy 104:135–145. https://doi.org/10.1016/j.marpol.2019.01.018

Emery TJ, Noriega R, Williams AJ, Larcombe J (2019b) Measuring congruence between electronic monitoring and logbook data in Australian Commonwealth longline and gillnet fisheries. Ocean & Coastal Management 168:307–321. https://doi.org/10.1016/j.ocecoaman.2018.11.003

Ewell C, Hocevar J, Mitchell E, et al (2020) An evaluation of Regional Fisheries Management Organization at-sea compliance monitoring and observer programs. Marine Policy 115:103842. https://doi.org/10.1016/j.marpol.2020.103842

Fangel K, Aas Ø, Vølstad JH, et al (2015) Assessing incidental bycatch of seabirds in Norwegian coastal commercial fisheries: Empirical and methodological lessons. Global Ecology and Conservation 4:127–136. https://doi.org/10.1016/j.gecco.2015.06.001

FAO Fisheries Department (2001) International Plan of Action to Prevent, Deter and Eliminate Illegal, Unreported and Unregulated Fishing. https://www.fao.org/3/y1224e/y1224e.pdf. Accessed 20th April 2022

Faunce CH (2011) A comparison between industry and observer catch compositions within the Gulf of Alaska rockfish fishery. ICES Journal of Marine Science 68:1769–1777. https://doi.org/10.1093/icesjms/fsr093

Fernandes AC, Oroszlányová M, Silva C, et al (2021) Investigating the representativeness of onboard sampling trips and estimation of discards based on clustering. Fisheries Research 234:105778. https://doi.org/10.1016/j.fishres.2020.105778

Figueiredo I, Maia C, Lagarto N, Serra-Pereira B (2020) Bycatch estimation of Rajiformes in multispecies and multigear fisheries. Fisheries Research 232:105727. https://doi.org/10.1016/j.fishres.2020.105727

Figus E, Criddle KR (2019) Comparing self-reported incidental catch among fishermen targeting Pacific halibut and a fishery independent survey. Marine Policy 100:371–381. https://doi.org/10.1016/j.marpol.2018.12.008

Fischer SH, De Oliveira JAA, Mumford JD, Kell LT (2021) Application of explicit precautionary principles in data-limited fisheries management. ICES Journal of Marine Science 78:2931–2942. https://doi.org/10.1093/icesjms/fsab169

Fiskeridirektøren (2022) Nasjonal strategisk risikovurdering (NSRV) for 2022. https://www.fiskeridir.no/Yrkesfiske/Nyheter/2022/nasjonal-strategisk-risikovurdering-for-2022. Accessed 16th April 2022

Ford E, Stewart BD (2021) Searching for a bridge over troubled waters: An exploratory analysis of trust in United Kingdom fisheries management. Marine Policy 132:104686. https://doi.org/10.1016/j.marpol.2021.104686

Fox DS, Starr RM (1996) Comparison of commercial fishery and research catch data. Canadian Journal of Fisheries and Aquatic Sciences 53:2681–2694. https://doi.org/10.1139/f96-230

Funtowicz S, Ravetz JR (1990) Uncertainty and Quality in Science for Policy, 1st edn. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-0621-1

Garcia S, Cochrane K (2005) Ecosystem approach to fisheries: a review of implementation guidelines. ICES Journal of Marine Science 62:311–318. https://doi.org/10.1016/j.icesjms.2004.12.003

Garcia SM, Kolding J, Rice J, et al (2012) Reconsidering the Consequences of Selective Fisheries. Science 335:1045–1047. https://doi.org/10.1126/science.1214594

Gezelius SS (2006) Monitoring fishing mortality: Compliance in Norwegian offshore fisheries. Marine Policy 30:462–469. https://doi.org/10.1016/j.marpol.2005.06.004

Gezelius SS (2002) Do norms count? State regulation and compliance in a Norwegian fishing community. Acta Sociologica 45:305–314. https://doi.org/10.1177/000169930204500404

Giacomarra M, Crescimanno M, Vrontis D, et al (2021) The ability of fish ecolabels to promote a change in the sustainability awareness. Marine Policy 123:104292. https://doi.org/10.1016/j.marpol.2020.104292

Gilman E, Perez Roda A, Huntington T, et al (2020) Benchmarking global fisheries discards. Scientific Reports 10:1–8. https://doi.org/10.1038/s41598-020-71021-x

Gray CA, Kennelly SJ (2018) Bycatches of endangered, threatened and protected species in marine fisheries. Reviews in Fish Biology and Fisheries 28:521–541. https://doi.org/10.1007/s11160-018-9520-7

Gullestad P, Abotnes AM, Bakke G, et al (2017) Towards ecosystem-based fisheries management in Norway – Practical tools for keeping track of relevant issues and prioritising management efforts. Marine Policy 77:104–110. https://doi.org/10.1016/j.marpol.2016.11.032

Gullestad P, Aglen A, Bjordal Å, et al (2014) Changing attitudes 1970–2012: evolution of the Norwegian management framework to prevent overfishing and to secure long-term sustainability. ICES Journal of Marine Science 71:173–182. https://doi.org/10.1093/icesjms/fst094

Gullestad P, Blom G, Bakke G, Bogstad B (2015) The "Discard Ban Package": Experiences in efforts to improve the exploitation patterns in Norwegian fisheries. Marine Policy 54:1–9. https://doi.org/10.1016/j.marpol.2014.09.025

Hareide N-R, Garnes G (2002) Misreporting – Results from Norwegian investigations on coastal gillnet fisheries, 2000. Working document no. 14 in the Report of the Arctic Fisheries Working Group, ICES CM 2002/ACFM:18

Harrison XA, Donaldson L, Correa-Cano ME, et al (2018) A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ 6:e4794. https://doi.org/10.7717/peerj.4794

Hatlebrekke HH (2021) A descriptive analysis of data from the Norwegian Coastal Reference Fleet and comparison with catch-per-unit-effort data from research surveys. Masters thesis, University of Bergen. Bergen Open Research Archive https://hdl.handle.net/11250/2756353. Accessed 20th April 2022

Hatlebrekke HH, Gundersen S, Nedreaas K, et al (2021) The Coastal Reference Fleet 2007-2019 - Fleet composition, fishing effort and contributions to science. Rapport fra havforskningen 2021-52 https://hdl.handle.net/11250/2837761. Accessed 20th April 2022

Hazen E, Scales KL, Maxwell SM, et al (2018) A dynamic ocean management tool to reduce bycatch and support sustainable fisheries. Science Advances 4:eaar3001. https://doi.org/10.1126/sciadv.aar3001

Helle K, Pennington M (2004) Survey design considerations for estimating the length composition of the commercial catch of some deep-water species in the northeast Atlantic. Fisheries Research 70:55–60. https://doi.org/10.1016/j.fishres.2004.06.011

Hoare D, Graham N, Schon PJ (2011) The Irish Sea data-enhancement project: comparison of self-sampling and national data-collection programmes—results and experiences. ICES Journal of Marine Science 68:1778–1784. https://doi.org/10.1093/icesjms/fsr100

Hønneland G (2000) Compliance in the Barents Sea Fisheries. How fishermen account for conformity with rules. Marine Policy 24:11–19. https://doi.org/10.1016/S0308-597X(98)00058-X

Hylen A (1965) On the estimation of cod and haddock discarded by trawlers using different chafers. Gadoid Fish Committee, ICES CM Documents;1965/167. http://hdl.handle.net/11250/101374. Accessed 20th April 2022

Hylen A (1967) Discarding of fish in North-East Atlantic. ICES CM Documents;1967/F:35. http://hdl.handle.net/11250/101199. Accessed 20th April 2022

Hylen A, Jacobsen JA (1987) Estimation of cod taken as by-catch in the Norwegian fishery for shrimp north of 69 °N. ICES CM 1987/G:34. https://www.ices.dk/sites /pub/CM%20Doccuments/1987/G/1987_G34.pdf. Accessed 20th April 2022

Hylen A, Smedstad OM (1974) Observations from the Barents Sea in spring 1973 on the discarding of cod and haddock caught in bottom and midwater trawls fitted with double cod ends. ICES CM Documents;1974/F:45 http://hdl.handle.net/11250/101766. Accessed 20th April 2022

ICES (2000) Report of the study group on discard and by-catch information. ICES, ICES CM 2000/ACFM:11. Copenhagen, Denmark https://www.ices.dk/community/ Documents/PGCCDBS/SGDBI01.pdf. Accessed 20th April 2022

ICES (2003) Workshop on discard sampling methodology and raising procedures. ICES Planning Group on Commercial Catch, Discards and Biological Sampling, Chalottenlund, Denmark 2-4 September 2003. https://www.ices.dk/community/ Documents/PGCCDBS/Discard_workshop%20_2003.pdf. Accessed 20th April 2022

ICES (2007a) Report on the workshop on discard raising procedures. ICES WKDRP CM 2007/ACFM:06 Ref RMC PGCCDBS. San Sebastian, Spain. https://www.ices.dk/sites/pub/Publication%20Reports/Expert%20Group%20Report/ac fm/PGCCDBS/WKDRP07.pdf. Accessed 20th April 2022

ICES (2007b) Report of the Workshop on Using Fishers to Sample Catches (WKUFS). CM 2007/ACFM:24. https://www.ices.dk/sites/pub/Publication%20Reports/Expert% 20Group%20Report/acfm/PGCCDBS/WKUFS_2007.pdf. Accessed 20th April 2022

ICES (2021) Arctic Fisheries Working Group (AFWG). ICES Scientific Reports. 3:58. 817 pp. https://doi.org/10.17895/ices.pub.8196

Jacobsen RB, Wilson DCK, Ramirez-Monsalve P (2012) Empowerment and regulation - dilemmas in participatory fisheries science. Fish and Fisheries 13:291–302. https://doi.org/10.1111/j.1467-2979.2011.00434.x

James KM, Campell N, Viðarsson JR, et al (2019) Tools and Technologies for the Monitoring, Control and Surveillance of Unwanted Catches. In: Uhlmann SS, Ulrich C, Kennelly SJ (eds) The European Landing Obligation: Reducing Discards in Complex, Multi-Species and Multi-Jurisdictional Fisheries. Springer, Cham, pp 363–382. https://doi.org/10.1007/978-3-030-03308-8

Janssen PHM, Petersen AC, van der Sluijs JP, et al (2005) A guidance for assessing and communicating uncertainties. Water Science and Technology 52:125–131. https://doi.org/10.2166/wst.2005.0160

Johnsen JP, Eliasen S (2011) Solving complex fisheries management problems: What the EU can learn from the Nordic experiences of reduction of discards. Marine Policy 35:130–139. https://doi.org/10.1016/j.marpol.2010.08.011

Karp WA, Breen M, Borges L, et al (2019) Strategies Used Throughout the World to Manage Fisheries Discards – Lessons for Implementation of the EU Landing Obligation. In: Uhlmann SS, Ulrich C, Kennelly SJ (eds) The European Landing Obligation: Reducing Discards in Complex, Multi-Species and Multi-Jurisdictional Fisheries. Springer, Cham, pp 3–26. https://doi.org/10.1007/978-3-030-03308-8

Kelleher K (2005) Discards in the world's fisheries. An update. FAO Fisheries Technical Paper. 470. Rome, Italy. https://www.fao.org/3/y5936e/y5936e00.htm. Accessed 20th April 2022

Kennelly SJ (2020) Bycatch Beknown: Methodology for jurisdictional reporting of fisheries discards – Using Australia as a case study. Fish and Fisheries 21:1046–1066. https://doi.org/10.1111/faf.12494

Kennelly SJ (2018) Developing a National Bycatch Reporting System. Final FRDC Report. FRDC Project No 2015/208. https://www.frdc.com.au/sites/default/files/products/2015-208-DLD.pdf. Accessed 20th April 2022

Kotwicki S, Ono K (2019) The effect of random and density-dependent variation in sampling efficiency on variance of abundance estimates from fishery surveys. Fish and Fisheries 20:760–774. https://doi.org/10.1111/faf.12375

Kraan M, Uhlmann S, Steenbergen J, et al (2013) The optimal process of self-sampling in fisheries: lessons learned in the Netherlands. Journal of Fish Biology 83:963–973. https://doi.org/10.1111/jfb.12192

Kvamme C, Ajiad A, Aglen A, Nedreaas K (2007) Estimated bycatch of haddock (*Melanogrammus aeglefinus*) and Greenland halibut (*Reinhardtius hippoglossoides*) in the Barents Sea shrimp fishery during 2000-2005. NAFO SCR Doc. 07/87. Serial No. N5473

Le Bot T, Lescroël A, Grémillet D (2018) A toolkit to study seabird–fishery interactions. ICES Journal of Marine Science 75:1513–1525. https://doi.org/10.1093/icesjms/fsy038

Lehuta S, Girardin R, Mahévas S, et al (2016) Reconciling complex system models and fisheries advice: Practical examples and leads. Aquatic Living Resources 29:208. https://doi.org/10.1051/alr/2016022

Li B, Chen Y, Reardon K, Wilson C (2019) A simulation-based evaluation of sea sampling design for American lobster in the Gulf of Maine. Fisheries Research 216:145–154. https://doi.org/10.1016/j.fishres.2019.04.012

Lohr SL (2021) Sampling: Design and Analysis, 3rd edn. Chapman and Hall/CRC. https://doi.org/10.1201/9780429298899

Lordan C, Ó Cuaig M, Graham N, Rihan D (2011) The ups and downs of working with industry to collect fishery-dependent data: the Irish experience. ICES Journal of Marine Science 68:1670–1678. https://doi.org/10.1093/icesjms/fsr115

Mackinson S (2022) The fall and rise of industry participation in fisheries science – a European story. ICES Journal of Marine Science, fsac041. https://doi.org/10.1093/icesjms/fsac041

Mackinson S, Middleton DAJ (2018) Evolving the ecosystem approach in European fisheries: Transferable lessons from New Zealand's experience in strengthening stakeholder involvement. Marine Policy 90:194–202. https://doi.org/10.1016/j.marpol.2017.12.001

Mangi SC, Dolder PJ, Catchpole TL, et al (2013) Approaches to fully documented fisheries: practical issues and stakeholder perceptions. Fish and Fisheries 16:426–452. https://doi.org/10.1111/faf.12065

Mangi SC, Kupschus S, Mackinson S, et al (2018) Progress in designing and delivering effective fishing industry-science data collection in the UK. Fish and Fisheries 19:622–642. https://doi.org/10.1111/faf.12279

Mangi SC, Smith S, Catchpole TL (2016) Assessing the capability and willingness of skippers towards fishing industry-led data collection. Ocean & Coastal Management 134:11–19. https://doi.org/10.1016/j.ocecoaman.2016.09.027

Martin TG, Wintle BA, Rhodes JR, et al (2005) Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecology Letters 8:1235–1246. https://doi.org/10.1111/j.1461-0248.2005.00826.x

McBride MM, Fotland Å (1996) Estimation of unreported catch in a commercial trawl fishery. Journal of the Northwest Atlantic Fisheries Society 18:31–41

Mendo T, Mendo J, Ransijn JM, et al (2022) Assessing discards in an illegal small-scale fishery using fisher-led reporting. Reviews in Fish Biology and Fisheries. https://doi.org/10.1007/s11160-022-09708-9

Mion M, Piras C, Fortibuoni T, et al (2015) Collection and validation of self-sampled e-logbook data in a Mediterranean demersal trawl fishery. Regional Studies in Marine Science 2:76–86. https://doi.org/10.1016/j.rsma.2015.08.009

Moan A (2016) Bycatch of harbour porpoise, harbour seal and grey seal in Norwegian gillnet fisheries. Masters thesis, University of Oslo. DUO Research Archive. http://urn.nb.no/URN:NBN:no-56897. Accessed 20th April 2022

Moan A, Skern-mauritzen M, Vølstad JH, Bjørge A (2020) Assessing the impact of fisheries-related mortality of harbour porpoise (*Phocoena phocoena*) caused by incidental bycatch in the dynamic Norwegian gillnet fisheries. ICES Journal of Marine Science 77:3039–3049. https://doi.org/10.1093/icesjms/fsaa186

Moore JE, Heinemann D, Francis TB, et al (2021) Estimating Bycatch Mortality for Marine Mammals: Concepts and Best Practices. Frontiers in Marine Science 8:752356. https://doi.org/10.3389/fmars.2021.752356

Nærings- og Fiskeridepartementet (2004) Kommisjonen for tiltak mot utkast av fisk - rapport med anbefalinger. https://www.regjeringen.no/no/dokumenter/utkastkommisjonens-rapport/id106098/. Accessed 20th April 2022

Nakagawa S, Cuthill IC (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. Biological Reviews 82:591–605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

Nedreaas K, Iversen S, Grethe K (2015) Preliminary estimates of total removals by the Norwegian marine fisheries, 1950-2010. Working Paper #2015-94, Fisheries Centre, University of British Columbia, Vancouver, Canada. http://www.seaaroundus.org/doc/publications/wp/2015/Nedreaas-et-al-Norway.pdf. Accessed 20th April 2022

NMFS (2011) U.S. National Bycatch Report. U.S. Department of Commerce, NOAA Technical Memo NMFS-F/SPO-117E, 508 p. https://www.fisheries.noaa.gov/resource/document/national-bycatch-report. Accessed 20th April 2022

Northridge S, Coram A, Kingston A, Crawford R (2017) Disentangling the causes of protected-species bycatch in gillnet fisheries. Conservation Biology 31:686–695. https://doi.org/10.1111/cobi.12741

NOU (2019) Framtidens fiskerikontroll. Departementenes sikkerhets- og serviceorganisasjon 2019:21 Oslo. https://www.regjeringen.no/no/dokumenter/nou-2019-21/id2680187. Accessed 20th April 2022

O'Neill FG, Feekings J, Fryer RJ, et al (2019) Discard avoidance by improving fishing gear selectivity: helping the fishing industry help itself. In: Uhlmann SS, Ulrich C, Kennelly SJ (eds) The European Landing Obligation: Reducing Discards in Complex, Multi-Species and Multi-Jurisdictional Fisheries. Springer, Cham, pp 279–296. https://doi.org/10.1007/978-3-030-03308-8

Ono K, Punt AE, Hilborn R (2015) Think outside the grids: An objective approach to define spatial strata for catch and effort analysis. Fisheries Research 170:89–101. https://doi.org/10.1016/j.fishres.2015.05.021

Österblom H, Sissenwine M, Symes D, et al (2011) Incentives, social–ecological feedbacks and European fisheries. Marine Policy 35:568–574. https://doi.org/10.1016/j.marpol.2011.01.018

Pálsson ÓK (2003) A length-based analysis of haddock discards in Icelandic fisheries. Fisheries Research 59:437–446. https://doi.org/10.1016/S0165-7836(02)00017-6

Pennington M (1996) Estimating the mean and variance from highly skewed marine data. Biometrics 39:281–286. https://doi.org/10.2307/2530830

Pennington M, Helle K (2011) Evaluation of the design and efficiency of the Norwegian self-sampling purse-seine reference fleet. ICES Journal of Marine Science 68:1764–1768. https://doi.org/10.1093/icesjms/fsr018

Pérez Roda MA, Gilman E, Huntington T, et al (2019) A third assessment of global marine fisheries discards. FAO Fisheries and Aquaculture Technical Paper No. 633. Rome, FAO. 78 pp. https://www.fao.org/3/CA2905EN/ca2905en.pdf. Accessed 20th April 2022

Perretti CT, Deroba JJ, Legault CM (2020) Simulation testing methods for estimating misreported catch in a state-space stock assessment model. ICES Journal of Marine Science 77:911–920. https://doi.org/10.1093/icesjms/fsaa034

Pikitch EK, Santora C, Babcock EA, et al (2004) Ecosystem-based fishery management. Science 305:346–347. https://doi.org/10.1126/science.1098222

Pitcher TJ, Kalikoski D, Pramod G, Short K (2009) Not honouring the code. Nature 457:658–659. https://doi.org/10.1038/457658a

Pitcher TJ, Watson R, Forrest R, et al (2002) Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish and Fisheries 3:317–339. https://doi.org/10.1046/j.1467-2979.2002.00093.x

Rago PJ, Wigley SE, Fogarty MJ (2005) NEFSC Bycatch Estimation Methodology: Allocation, Precision, and Accuracy. U.S. Dep. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 05-09. https://repository.library.noaa.gov/view/noaa/3455. Accessed 20th April 2022

Rayner S (2012) Uncomfortable knowledge: the social construction of ignorance in science and environmental policy discourses. Economy and Society 41:107–125. https://doi.org/10.1080/03085147.2011.637335

Roman S, Jacobson N, Cadrin SX (2011) Assessing the reliability of fisher self-sampling programs. North American Journal of Fisheries Management 31:165–175. https://doi.org/10.1080/02755947.2011.562798

Rosenberg AA (2007) Fishing for certainty. Nature 449:989–989. https://doi.org/10.1038/449989a

Rudd MB, Branch TA (2017) Does unreported catch lead to overfishing? Fish and Fisheries 18:313–323. https://doi.org/10.1111/faf.12181

Saltelli A (2019) A short comment on statistical versus mathematical modelling. Nature Communications 10:3870. https://doi.org/10.1038/s41467-019-11865-8

Saltelli A, Bammer G, Bruno I, et al (2020) Five ways to ensure that models serve society: a manifesto. Nature 582:482–484. https://doi.org/10.1038/d41586-020-01812-9

Sandin P (1999) Dimensions of the Precautionary Principle. Human and Ecological Risk Assessment: An International Journal 5:889–907. https://doi.org/10.1080/10807039991289185

Skern-Mauritzen M, Ottersen G, Handegard NO, et al (2016) Ecosystem processes are rarely included in tactical fisheries management. Fish and Fisheries 17:165–175. https://doi.org/10.1111/faf.12111

Sohns A, Hickey GM, Temby O (2022) Exploring the potential impacts of machine learning on trust in fishery management. Fish and Fisheries. https://doi.org/10.1111/faf.12658

Starr P (2010) Fisher-Collected Sampling Data: Lessons from the New Zealand Experience. Marine and Coastal Fisheries 2:47–59. https://doi.org/10.1577/c08-030.1

STECF (2013) Landing obligation in EU fisheries (STECF-13-23). Publications Office of the European Union, Luxembourg, EUR 26330 EN, JRC 86112, 115 pp. https://doi.org/10.2788/37460

Stock BC, Ward EJ, Eguchi T, et al (2020) Comparing predictions of fisheries bycatch using multiple spatiotemporal species distribution model frameworks. Canadian Journal of Fisheries and Aquatic Sciences 77:146–163. https://doi.org/10.1139/cjfas-2018-0281

Stock BC, Ward EJ, Thorson JT, et al (2019) The utility of spatial model-based estimators of unobserved bycatch. ICES Journal of Marine Science 76:255–267. https://doi.org/10.1093/icesjms/fsy153

Straan K-B, Vader W, Barrett R (1991) Auk mortality in fishing nets in north Norway. Seabird 13:22–29

Stratoudakis Y, Fryer RJ, Cook RM, Pierce GJ (1999) Fish discarded from Scottish demersal vessels: Estimators of total discards and annual estimates for targeted gadoids. ICES Journal of Marine Science 56:592–605. https://doi.org/10.1006/jmsc.1999.0500

Sturludottir E, Sigurdsson GM, Stefansson G (2019) Evaluation of a length-based method to estimate discard rate and the effect of sampling size. arXiv preprint: not peer reviewed: https://doi.org/10.48550/arXiv.1902.10579

Suuronen P, Gilman E (2020) Monitoring and managing fisheries discards: New technologies and approaches. Marine Policy 116:103554. https://doi.org/10.1016/j.marpol.2019.103554

Svorken M, Hermansen Ø (2014) Urapportert fiske i torskefiskeriene - Resultater fra spørreundersøkelse om juks. Nofrima report number 26/2014. http://hdl.handle.net/11250/283688. Accessed 20th April 2022

Thorson JT (2019) Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research 210:143–161. https://doi.org/10.1016/j.fishres.2018.10.013

Tirrell A (2017) Sociocultural institutions in Norwegian fisheries management. Marine Policy 77:37–43. https://doi.org/10.1016/j.marpol.2016.12.002

Trochta JT, Pons M, Rudd MB, et al (2018) Ecosystem-based fisheries management: Perception on definitions, implementations, and aspirations. PLoS ONE 13:e0190467. https://doi.org/10.1371/journal.pone.0190467

Valdemarsen JW, Nakken O (2002) Utkast i norske fiskerier. Workshop om utkast i nordiske fiskerier, Sophienberg Slot 18-20 november 2002

van der Sluijs JP, Craye M, Funtowicz S, et al (2005) Combining Quantitative and Qualitative Measures of Uncertainty in Model-Based Environmental Assessment: The NUSAP System. Risk Analysis 25:481–492. https://doi.org/10.1111/j.1539-6924.2005.00604.x

van der Sluijs JP, Petersen AC, Janssen PHM, et al (2008) Exploring the quality of evidence for complex and contested policy decisions. Environmental Research Letters 3:024008. https://doi.org/10.1088/1748-9326/3/2/024008

Vigneau J (2006) Raising procedures for discards: Sampling theory. ICES CM 2006/K :16, ICES, Copenhagen, Denmark

Wade PR, Long KJ, Francis TB, et al (2021) Best Practices for Assessing and Managing Bycatch of Marine Mammals. Frontiers in Marine Science 8:1566. https://doi.org/10.3389/fmars.2021.757330

Walls RHL, Dulvy NK (2021) Tracking the rising extinction risk of sharks and rays in the Northeast Atlantic Ocean and Mediterranean Sea. Scientific Reports 11:15397. https://doi.org/10.1038/s41598-021-94632-4

Walsh WA, Ito RY, Kawamoto KE, McCracken M (2005) Analysis of logbook accuracy for blue marlin (*Makaira nigricans*) in the Hawaii-based longline fishery with a generalized additive model and commercial sales data. Fisheries Research 75:175–192. https://doi.org/10.1016/j.fishres.2004.11.007

Wigley SE, Asci S, Benjamin S, et al (2021) Standardized bycatch reporting methodology 3-year review report - 2020: Reviewing SBRM years 2018, 2019, and

2020. NOAA Technical Memorandum NMFS-NE-266. https://doi.org/10.25923/d4xb-4183

Wigley SE, Palmer MC, Blaylock J, Rago PJ (2008) A Brief Description of the Discard Estimation for the National Bycatch Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 08-02. https://repository.library.noaa.gov/view/noaa/3531. Accessed 20th April 2022

Williams T, Gundersen S (2021) Havforskningsinstituttets Referanseflåte - Årsrapport 2020/2021. Institute of Marine Research, Bergen, Norway.

Williams T, Gundersen S, Clegg TL (2018) The Norwegian Reference Fleet - the stakeholder's perception of the collaboration. In: Proceedings of the 9th International Fisheries Observer and Monitoring Conference. Vigo, Spain. https://ifomcvigo.com/wp-content/uploads/2018/08/proceedings-9th-ifomc1.pdf. Accessed 20th April 2022

Yan Y, Cantoni E, Field C, et al (2021) Spatiotemporal modeling of bycatch data: methods and a practical guide through a case study in a Canadian Arctic fishery. Canadian Journal of Fisheries and Aquatic Sciences 79:148–158. https://doi.org/10.1139/cjfas-2020-0267

Zeller D, Cashion T, Palomares M, Pauly D (2017) Global marine fisheries discards: A synthesis of reconstructed data. Fish and Fisheries 19:30–39. https://doi.org/10.1111/faf.12233

Glossary

Accuracy	The combination of bias and precision.
Bias	A systematic error in the sampling, measurement, or estimation procedure that results in a statistic being consistently larger (or smaller) than the population characteristic that it estimates.
Bycatch	The catch of nontarget animals, which can either be landed or discarded. This includes juveniles and undersized specimens of the target species.
Discard ban	A legal requirement that certain catches of fish are not allowed to be discarded at sea. Exemptions may occur for species or circumstances. Often used synonymously with *landing obligation* if both are implemented simultaneously.
Discard rate	The proportion (or percentage) of the total catch (numbers or weight) that is discarded.
Discards	The portion of animals in the total catch which is thrown away or dumped at sea before landing for whatever reason.
Fishery	A combination of spatial area, fishing gear, and target species. A temporal element may also be included if discrete. Note that these factors may be difficult to define, such as in multi-species fisheries.
Fishing operation	Typically defined as the period between the setting and hauling of a single continuous piece of fishing gear.
High-grading	The act of discarding individual fish with a lower commercial value to make room for catches with higher commercial value when space or quota is limited.
ICES	The International Council for the Exploration of the Sea.
IMR	Norwegian Institute of Marine Research.
Incidental catches	Bycatches with no commercial value that the vessel had no intention of catching (e.g. seabirds, marine mammals).
IUU fishing	Illegal, unreported, and unregulated fishing.
Landing obligation	A legal requirement to land and report all catches. Exemptions may occur for non-quota species (see also *discard ban*).
Landings	Catches retained on board and landed upon returning to port.
Precision	The amount of variation among estimates from repeated samples.

Target species	One or more species which a fisher intends to capture. In mixed fisheries, it is sometimes difficult to define the target species as it depends on many unobserved factors or includes opportunistic factors.
Total catches	All biological material retained by the fishing gear and brought on board the vessel. May also be used in the context of a single species.
Unmandated catches	Catches for which there is no legal requirement to explicitly report upon landing.
Unreported catches	Catches that are not reported explicitly in official statistics. They comprise of unmandated catches, illegal catches, and discards.
Unwanted catches	Bycatches of threatened species and species without economic interest to the fisher, as well as species for which a particular fisher or fleet does not hold a quota or fishing right (Gullestad et al. 2015).

Appendix A: Unreported catches in the Barents Sea trawl fishery

The estimation of unreported catches in the Barents Sea trawl fishery was done on a limited number of 30 species due to data inaccuracies. For species processed in the on-board factory, the retained weight was entered instead of the total weight, therefore omitting the unreported portion from the value. Whilst further investigation could determine which species were retained on the level of individual hauls, the analysis here uses the short-term solution of only estimating species which were observed by the Norwegian Reference Fleet but do not appear in the sales notes for any given year. Therefore, total catches estimated here are entirely unreported.

The Barents Sea trawl fishery is described in **Paper II**. Total catches were estimated using the methodology described in **Paper III**, but only applying the cluster unit and cluster ratio estimator (using trawl duration as the measure of effort).

A preliminary analysis found that the correlation between total catches and trawl duration was poor across all species, with Pearson's correlation coefficients ranging from -0.10–0.16. Such poor relationships are expected for the rarer species included in this analysis (0.40–38.64 % encounter rate). Trawl vessels participating in the Norwegian Reference Fleet tended to be the most active vessels in the fishery (Figure A1). This reflects the same pattern in representativeness observed in the longline fishery (**Paper III**). Between 2012 and 2014, sampled vessels had on average shorter trawl durations than unobserved vessels in the fishery. However, this trend disappears in later years where Norwegian Reference Fleet vessels had trawl durations typical of the fishery. As both the unit and ratio estimator have weak assumptions, both are applied.

Comparisons of annual total catches of species using the cluster unit and cluster ratio estimators (Figure A2) found that only one third of estimates were similar (difference < 20 %). However, the choice of estimator had little effect on precision, with coefficient of variation being similar for the majority of estimates. All estimates of total catches are provided in Table A1.

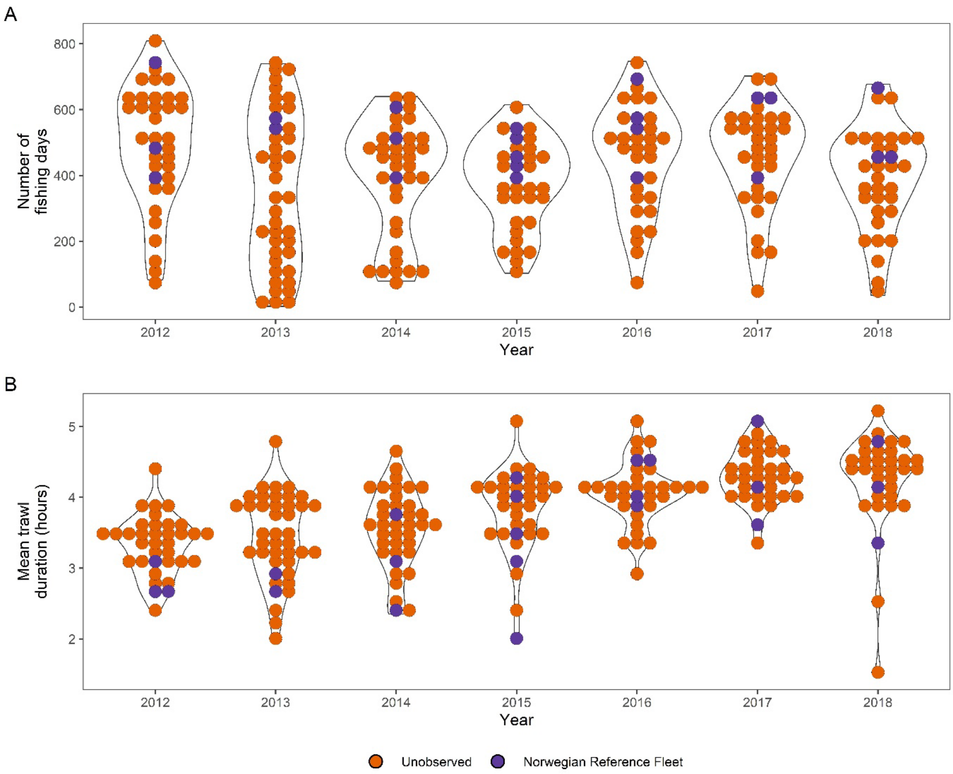


Figure A1. Representativeness of sampled fishing effort in the Barents Sea trawl fishery. (A) Number of fishing days per vessels; (B) Mean trawl duration (hours), Each point represents one vessel.

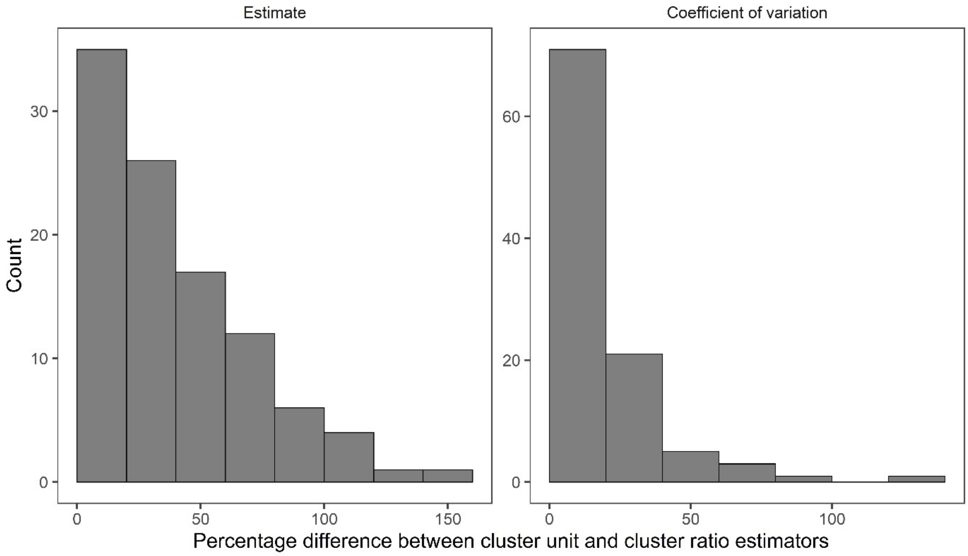


Figure A2. Difference in estimate and variation between cluster unit and cluster ratio estimators.

Table A1. Estimates (and 95 % confidence interval) of total catches for a limited selection of species. ER = encounter rate.

Year	Latin name	ER (%)	Estimated total catches (kg)	
			Cluster ratio estimator	Cluster unit estimator
2012	*Chimaera monstrosa*	2.0	1224 (0–24762)	1829 (0–63819)
	Clupea harengus	4.1	672 (306–9932)	753 (380–14992)
	Cyclopterus lumpus	7.3	14353 (8180–46894)	16558 (10664–49250)
	Eutrigla gurnardus	0.4	107 (0–42463)	97 (0–30023)
	Glyptocephalus cynoglossus	1.6	615 (64–1504)	743 (130–2144)
	Lepidorhombus whiffiagonis	1.6	878 (0–2376)	1180 (0–4641)
	Lycodes esmarkii	2.4	2368 (13–4997)	3299 (75–8469)
	Macrourus berglax	0.4	889 (0–8750)	642 (0–16392)
	Mallotus villosus	0.4	6 (0–15)	38 (0–107)
	Micromesistius poutassou	3.6	1309 (276–4195)	1415 (365–22592)
	Microstomus kitt	4.1	1660 (515–5062)	2081 (675–4787)
	Molva dypterygia	0.4	373 (0–1067)	397 (0–1701)
	Myoxocephalus scorpius	0.4	16 (0–39)	7 (0–26)
	Phycis blennoides	1.6	547 (0–9751)	560 (0–15907)
	Salmo salar	0.4	999 (0–3105)	804 (0–1883)
	Scomber scombrus	1.2	86 (0–235)	113 (0–269)
	Sebastes viviparus	18.6	54154 (32112–87994)	54406 (34447–76550)
	Trachurus trachurus	0.4	98 (0–471)	119 (0–339)
	Trisopterus esmarkii	1.6	199 (0–551)	223 (0–720)
2013	*Chimaera monstrosa*	1.9	1420 (341–23821)	2908 (651–61860)
	Clupea harengus	1.9	451 (222–9824)	883 (448–14651)
	Cyclopterus lumpus	23.3	41212 (12097–65822)	45127 (13451–76466)
	Eutrigla gurnardus	1.9	10951 (10–37601)	9151 (24–39108)
	Glyptocephalus cynoglossus	1.9	606 (160–1872)	664 (213–1759)
	Lepidorhombus whiffiagonis	1.0	508 (27–1720)	903 (88–4523)
	Lycodes esmarkii	3.0	172 (47–1242)	305 (83–1119)
	Micromesistius poutassou	1.0	740 (440–2320)	1516 (924–17215)
	Microstomus kitt	4.9	1363 (432–3249)	2378 (673–4706)
	Pleuronectes platessa	1.9	242 (0–812)	307 (0–839)
	Salmo salar	1.0	997 (1–5761)	1249 (1–4166)
	Sebastes viviparus	14.6	25188 (15398–46954)	41943 (26251–61976)
2014	*Chimaera monstrosa*	9.1	84633 (3168–132521)	220271 (8438–414425)
	Clupea harengus	4.6	319 (29–938)	510 (122–2492)
	Coryphaenoides rupestris	2.3	10633 (0–20881)	22478 (0–67433)
	Cyclopterus lumpus	38.6	67797 (37745–127281)	85337 (53166–124433)
	Eutrigla gurnardus	15.9	16971 (394–34721)	19363 (1202–42706)
	Glyptocephalus cynoglossus	6.8	1456 (255–3012)	2681 (307–5638)
	Lepidorhombus whiffiagonis	15.9	12376 (382–15785)	29010 (900–40304)
	Lycodes esmarkii	2.3	560 (65–1597)	861 (125–2719)
	Mallotus villosus	2.3	13 (0–39)	41 (0–122)
	Micromesistius poutassou	9.1	45646 (3620–104441)	97213 (7466–150024)
	Phycis blennoides	15.9	7449 (1004–17417)	12552 (3640–19247)
	Salmo salar	2.3	137 (0–534)	420 (0–1260)
	Scyliorhinus canicula	2.3	231 (0–760)	487 (0–1460)
	Sebastes viviparus	9.1	4883 (2740–9648)	15751 (8275–26935)
	Solea solea	2.3	205 (0–594)	240 (0–721)
	Squalus acanthias	2.3	2802 (0–4753)	7235 (0–14592)
2015	*Chimaera monstrosa*	3.3	6870 (481–91151)	27867 (1777–237343)
	Clupea harengus	1.1	113 (7–765)	222 (61–2182)
	Cyclopterus lumpus	19.8	50030 (21679–104476)	60408 (35791–97346)

Year	Latin name	ER (%)	Estimated total catches (kg)	
			Cluster ratio estimator	Cluster unit estimator
	Eutrigla gurnardus	1.1	1088 (0–22795)	1233 (1–21317)
	Lepidorhombus whiffiagonis	4.4	3437 (69–13530)	9946 (331–35168)
	Micromesistius poutassou	7.7	24069 (1557–105949)	16771 (4388–139288)
	Microstomus kitt	5.6	887 (55–1895)	1605 (270–4055)
	Phycis blennoides	4.4	2785 (286–8927)	7474 (1142–15387)
	Sebastes viviparus	9.9	6373 (1871–10899)	18464 (7540–32575)
2016	*Centroscymnus crepidater*	1.1	911 (0–2866)	1209 (0–3623)
	Chimaera monstrosa	8.6	39561 (2510–83308)	51298 (4433–162818)
	Clupea harengus	4.3	3506 (38–8522)	5435 (90–11823)
	Cyclopterus lumpus	25.8	48844 (22927–88187)	68059 (33423–119496)
	Eutrigla gurnardus	1.1	254 (1–790)	266 (6–1171)
	Glyptocephalus cynoglossus	1.1	104 (1–683)	217 (7–1026)
	Lepidorhombus whiffiagonis	1.1	678 (3–2973)	667 (16–5147)
	Mallotus villosus	1.1	27 (0–67)	55 (0–132)
	Micromesistius poutassou	11.8	24383 (1346–53256)	33606 (2242–75282)
	Phycis blennoides	8.6	16400 (5983–27881)	24665 (10452–44770)
	Sebastes viviparus	1.1	1173 (23–19393)	2140 (122–28866)
	Squalus acanthias	1.1	186 (0–563)	296 (0–966)
2017	*Clupea harengus*	2.0	1522 (98–69779)	3245 (194–8953)
	Cottidae	0.7	149 (0–476)	104 (0–364)
	Cottunculus microps	0.7	23 (0–65)	11 (0–49)
	Cyclopterus lumpus	29.1	83950 (56409–111556)	97872 (66659–147587)
	Eutrigla gurnardus	0.7	420 (1–1623)	335 (8–1390)
	Galeus melastomus	0.7	68 (0–273)	73 (0–276)
	Glyptocephalus cynoglossus	1.3	2278 (3–8456)	2126 (12–5751)
	Lepidorhombus whiffiagonis	2.7	368 (6–3339)	303 (30–2161)
	Lycodes esmarkii	1.3	257 (2–887)	150 (10–430)
	Macrourus berglax	0.7	4057 (0–15400)	7102 (0–21895)
	Micromesistius poutassou	9.3	61297 (579–108796)	64875 (1792–139083)
	Microstomus kitt	3.3	350 (37–1407)	289 (65–2392)
	Phycis blennoides	2.7	2709 (177–6046)	2137 (404–6454)
	Salmo salar	1.3	1609 (0–4807)	1376 (1–3436)
	Scomber scombrus	0.7	70 (0–283)	74 (0–285)
	Sebastes viviparus	7.3	9729 (308–31892)	6621 (619–25435)
2018	*Chimaera monstrosa*	4.0	33719 (1768–108947)	33187 (6762–94254)
	Clupea harengus	2.0	43213 (97–99893)	42510 (283–115355)
	Cottunculus microps	0. 7	15 (0–36)	20 (0–66)
	Cyclopterus lumpus	25.3	60938 (32391–86101)	82031 (47635–118063)
	Etmopterus spinax	1.3	715 (0–1646)	702 (0–1990)
	Eutrigla gurnardus	1.3	480 (73–1190)	707 (171–1793)
	Galeus melastomus	0.7	573 (0–1464)	562 (0–1805)
	Glyptocephalus cynoglossus	3.3	1051 (14–4654)	1149 (47–3082)
	Lepidorhombus whiffiagonis	3.3	2075 (459–4381)	3033 (740–6801)
	Lycodes esmarkii	2.7	510 (6–1194)	502 (21–1379)
	Mallotus villosus	0.7	13 (0–36)	41 (0–115)
	Micromesistius poutassou	11.3	83107 (23074–143958)	60030 (17543–115101)
	Microstomus kitt	4.0	3225 (456–5393)	4522 (977–8301)
	Molva dypterygia	0.7	26 (0–226)	38 (0–120)
	Phycis blennoides	6.0	32932 (4414–65225)	32281 (5612–71869)
	Pleuronectidae	1.3	1093 (0–3617)	1297 (0–4331)
	Sebastes viviparus	6.0	5381 (982–13681)	9593 (2440–20869)
	Solea solea	0.7	113 (0–886)	85 (0–216)

Appendix B: Model-based raising procedure key

Candidate models
M1: log(discards + 1) ~ [stratum components] + (1 | vessel)
M2: log(discards + 1) ~ offset(log(effort)) + [stratum components] + (1 | vessel)

n.b. Formulae written using lme4 syntax of the R programming language.
[stratum components] fitted as individual fixed effects (e.g. … year + area + quarter …)

MODEL-BASED RAISING PROCEDURE KEY

1. Do all strata have enough observations to ensure model convergence?
   Yes — go to 2
   No — Re-stratify samples in a simpler system; go to 2

2. Fit candidate models (M1 and M2)
   Go to 3

3. Did the models converge?
   Yes — Go to 5
   No — Go to 4

4. Are convergence issues caused by the random vessel effect?
   Yes — Drop random effect; go to 2
   No — Go to 1

5. Perform cross-validation to evaluate predictive accuracy
   One model better — Use best model; go to 6
   Both models equal — Use both models; go to 6

6. Do results have actionable consequences?
   Yes — Go to 7
   No — END

7. Perform full model validation (e.g. check residual patterns, zero-inflation, spatial or temporal correlation)
   Issues identified — Improve model using information theoretic approach
   No issues identified — END

Key adapted from ICES (2007a)

Appendix C: Papers

PAPER I

Rev Fish Biol Fisheries
https://doi.org/10.1007/s11160-020-09624-w

REVIEWS

Applying global best practices for estimating unreported catches in Norwegian fisheries under a discard ban

Thomas L. Clegg · Steven J. Kennelly · Geir Blom · Kjell Nedreaas

Received: 7 February 2020 / Accepted: 23 October 2020

Abstract In addition to their role as a fisheries management tool, discard bans can be effective in improving knowledge of total catches via the requirement to land and report all catches. This shifts the focus to understanding the scale of unreported catches in fisheries, rather than only on discards. However, the presence of a discard ban can cause problems with estimation process, as it involves the observation of illegal activities, and the complex sources of unreported catches require a different approach to estimation. The Norwegian discard ban was introduced in 1987 as part of a wider suite of regulatory measures to improve exploitation patterns in commercial fisheries, but a framework for the regular estimation of unreported catches has yet to be established and operationalised. Here, we aim to identify global best practices for estimating unreported catches under a discard ban and assess their applicability to Norwegian fisheries. We approach this in three steps: (1) defining the scope of an estimation, (2) data collection, and (3) the actual procedure for estimation. We discuss how each step can affect the quality of an estimate with regards to accuracy, precision, practical limitations and whether the estimate is fit for purpose. Finally, we provide a list of recommendations for future studies and identify key knowledge gaps and limitations regarding their application to Norwegian fisheries.

Keywords Bycatch · Discards · Self-sampling · Ecosystem approach · Fisheries management · Stock assessment

Introduction

Information about total removals by a fishery is vital to detect and manage impacts on stocks and ecosystems and so contribute to the long-term sustainability of the fishery. However, if this knowledge comes from reported catches, then it only represents the landed portion of catches (hereafter referred to as landings). That is, such data do not give a complete picture of total extractions because of discarding at sea and any catches that are misreported or not reported at all.

Many of today's stock assessments use reported catch statistics to estimate population abundance and fishing mortality which lead to management

T. L. Clegg (✉) · K. Nedreaas
Institute of Marine Research/Havforskningsinstituttet, PO Box 1870, Nordnes, 5817 Bergen, Norway
e-mail: tom.clegg@hi.no

S. J. Kennelly
IC Independent Consulting, Cronulla, NSW, Australia

G. Blom
Directorate of Fisheries, PO Box 185, 5804 Bergen, Norway

Published online: 24 December 2020

recommendations, so it is vital that all catches are accounted for. Inaccurate reporting can affect estimations for those assessments (Dickey-Collas et al. 2007; Rudd and Branch 2017) and have specific effects on outputs concerning undersized fishes, such as recruitment (Punt et al. 2006; Dickey-Collas et al. 2007). For non-commercial species, a lack of understanding about total catches will limit knowledge of a fishery's impact on the wider ecosystem, particularly on species of conservation importance (Gray and Kennelly 2018). Knowledge of such bycatches are also necessary for eco-labelling initiatives, such as Marine Stewardship Council certification. In addition to environmental impacts, discarding is also perceived as a waste of resources. Public ownership of wild fisheries resources exists up to the point of retention, so discarded fish are effectively in permanent public ownership (Gray and Kennelly 2018). Governments and managers therefore have an obligation to monitor and reduce this wastage in the public interest. Wasted resources also have the potential to become new market opportunities, improving utilisation and economics sustainability.

A discard ban (also referred to as a landing obligation) can be an effective tool towards accounting for all catches in a fishery, as all catches are supposed to be landed and reported. In a global review of discard ban strategies, Karp et al. (2019) concluded that the success of a discard ban depends largely on the ability to enforce it, coupled with the acceptance and compliance of stakeholders. They also noted that discard bans may introduce complications in gathering high quality data on catches and discards at sea, and so restrict the ability to verify the effectiveness of a ban. These limitations are evident in recent global estimations of discards by Pérez Roda et al. (2019) and Gilman et al. (2020), where discard rates for Norway and Iceland had to be assumed due to low data availability.

Norway first introduced a discard ban on cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) in 1987 to address declining stocks of these species in the Barents Sea. A suite of regulatory measures was also introduced alongside, collectively referred to as the 'Discard Ban Package' (see Gullestad et al. 2015 for full description). The measures included real-time closures, compensation for the landing of illegal catches, and development of gear selectivity, all of which aimed to remove incentives for discarding by encouraging the avoidance of unwanted catches. Over the following decades, the discard ban was extended to include more species such that now, under the Marine Resources Act 2008, there is an obligation to land and report all catches. Under the current legislation, there are still exemptions to the obligation,[1] which include any fish that are alive when discarded, as well as certain protected species that must be released back into the sea immediately regardless of if they are alive or dead, but these must still be recorded in the catch logbook even though they were not retained.

There have been no direct studies that quantified the impact of the Norwegian discard ban on discarding practices, either as it developed or in the ensuing years. Nedreaas et al. (2015) reconstructed total catches for numerous fisheries between 1950 and 2010, reporting a overall decrease in unreported catches after the introduction of the discard ban. Other estimates of discards and unreported catches in Norway (Dingsør 2001a; Valdemarsen and Nakken 2002) indicate low levels of discarding relative to the global average (Pérez Roda et al. 2019), whilst numerous studies have provided snapshot estimates for individual fisheries (e.g. Hylen and Jacobsen 1987; McBride and Fotland 1996; Dingsør 2001b; Breivik et al. 2017). The available estimates, both nationally and for individual fisheries, have been constrained by a lack of at-sea observations throughout time, focussing on shorter timescales and specific fisheries where data are available.

We therefore acknowledge that the Norwegian discard ban is difficult to enforce (Gezelius 2006; Gullestad et al. 2015; NOU 2019), and that the level of discarding in Norwegian waters is still relatively unknown (Gullestad et al. 2015; Nedreaas et al. 2015). The monitoring and management of unwanted catches is a core component of ecosystem-based fisheries management generally (Pikitch et al. 2004; Bellido et al. 2011), but for it to be effective, a better understanding is needed of the scale and causes of unreported catches, and the impacts on ecosystems. However, there is currently no system in place to provide regular estimates of unreported catches in Norway, which are necessary for stock assessments for

[1] As the list is updated intermittently under new legislation, the latest version can be found here: https://www.fiskeridir.no/English/Fisheries/Regulations.

commercial species and evidence-based management of bycatches.

In this review we aim to identify best practices used globally to estimate unreported bycatches and discards and determine if they can be applied to Norwegian fisheries under a discard ban. To achieve this, we have broken down the process into three stages: (1) defining the scope of a study, (2) data collection, then (3) the estimation procedure used. At each stage, we critically evaluate approaches from the literature to identify best practices, then assess the extent to which they can be applied to Norwegian fisheries, giving focus to the influence of the discard ban. A schematic diagram for this process is shown in Fig. 1, listing the themes addressed at each stage. Through this process, we identify best practice guidelines for estimations of unreported catches which are applicable to fisheries under a discard ban, whilst identifying knowledge gaps and limitations which should be addressed to improve estimations.

Defining the scope of estimating unreported catches

Defining the scope of a study beforehand helps to guide decisions on data collection and the estimation procedure. In addition, a well-defined scope will provide a firmer understanding of what inferences can be made once an estimation is obtained.

We have not considered some sources of unreported catches in this review due to them being out of scope. Marine recreational fishing has been shown to contribute substantially to total mortality in European fisheries, with evidence that removals from recreational fisheries can exceed commercial fishing in some cases (Radford et al. 2018). Therefore, recreational fisheries must be considered and accounted for in total removals. However, large differences in sampling approaches are needed to adequately address their unique dynamics (e.g. in fishing gear, catch and release practices) (National Research Council 2006), meaning that quantifying unreported catches in recreational fisheries is out of the scope of this review. Mortality of organisms that encounter fishing gear underwater but are not caught is not accounted for in total extractions, which can occur after escapement from gear before it is hauled, either through physical injury or stress (Veldhuizen et al. 2018). This is also applicable to habitat damage caused by fishing gears, particularly bottom trawls, which damage benthic community structures and habitats (Kaiser et al. 2006). Finally, mortality by abandoned fishing gears, known as ghost fishing, can continue to occur indefinitely. Whilst it can have large environmental impacts, it is often addressed in a different management framework (Gilman 2015) and requires a different sampling methodology to quantify mortality.

Terminology

The definitions used in this review are based on those of Kelleher (2005), with specific adaptations highlighted. A *fishery* is defined as a group of similar fishing gears targeting one or more species in a fishing area or zone. The *catch* (also referred to as 'gross catch') is all biological material retained by the fishing gear and brought on board the vessel. This differs from the definition given by Kelleher (2005) because estimating unaccounted mortality whilst the gear is underwater is not possible using on-board catch sampling methods considered here (see above). After the catch is brought on board and sorted, *landings* are the portion of the catch that is brought ashore. *Discards* are defined as that portion of animals in the catch which is thrown away or dumped at sea before landing for whatever reason. It does not include shells, corals, plants, or inorganic materials (sometimes considered a concern of environmental impact), nor processing waste such as offal and carcasses. Discards include slipping, an event typically associated with purse seine fisheries where catches are released before being brought on board. *Bycatch* is the catch of non-target animals, which can either be landed or discarded. This includes juveniles and undersized specimens of the target species. *Unreported catches* contain any catches that are not reported upon landing under a landing obligation. They can be separated into three general categories: unmandated catches, illegal catches, and discards (Pitcher et al. 2002). These are expanded upon in the next section.

The terms 'discard ban' and 'landing obligation' are used synonymously in many descriptions of discard reduction policies and are used as such in this review. However, they are also two distinct legal terms. By definition, a discard ban makes the act of discarding illegal, whilst a landing obligation creates the legal requirement to land and report all catches. This is seen in the history of Norwegian discard policy,

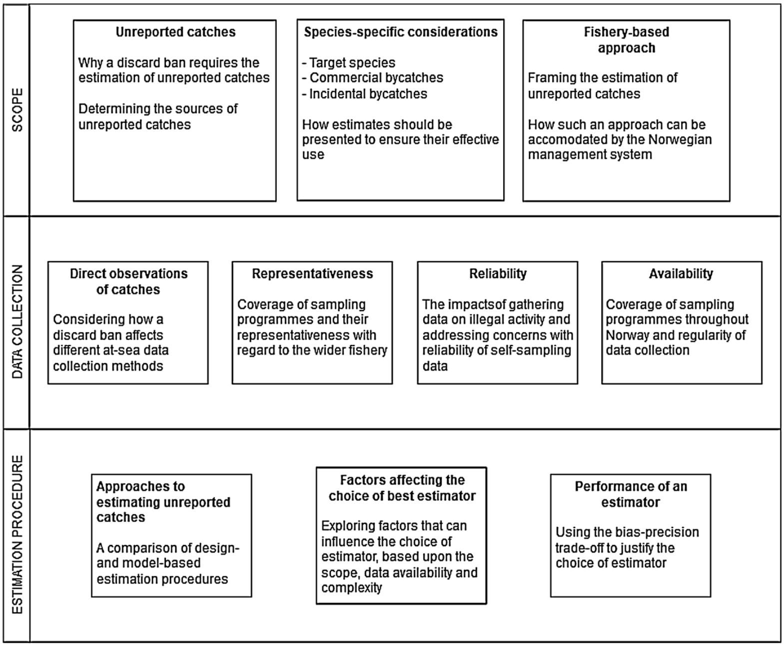


Fig. 1 Schematic diagram of the themes addressed in this review at each stage of the process for estimating unreported catches

where the act of discarding was banned in Norway in 1987, but it was only in 2009 that a "landing obligation" was introduced. In contrast, the reform of the EU common fisheries policy in 2014 introduced a discard ban and landing obligation simultaneously. Therefore, to assess the effectiveness of a discard ban, then total discards must be quantified. The same assessment for a landing obligation requires the quantification of unreported catches to assess the extent to which reported landings reflect total catches.

Unreported catches

Unmandated catches

Global reviews of discard ban policies by Borges et al. (2016) and Karp et al. (2019) found no examples where the discarding of all species is prohibited. Instead, discard bans have focussed on species with quota regulations, aiming to ensure that all catches count towards total catch allowances (e.g. the European Union and New Zealand), whilst others apply to a defined list of species that includes non-quota and non-commercial species, but are not exhaustive (e.g. Norway and Iceland).

While numerous discard bans have addressed the issue of mandatory reporting, there remain difficulties in the resolution of such reports. For some species groups, there can be no mandate to differentiate between individual species. This is particularly the case for elasmobranchs, for which there are substantial knowledge gaps in bycatch information worldwide (Oliver et al. 2015) due to difficulties in species identification and a general lack of reporting. Fishmeal production facilities on-board vessels cause similar problems if individual species contributions are not reported. Whilst all catches will have technically been accounted for in these situations, the lack of detail means they should still be classed as unreported catches for the purposes of estimation and management advice regarding individual species.

The Norwegian discard ban applies to all species in principal, but subsequent legislation has confined mandatory reporting to a list of 55 species or species groups. The overall resolution of species reporting is high across fisheries, but there are a small percentage of species reported to a higher taxonomic level. These are almost entirely elasmobranchs (especially skates and rays), for which species reporting is poor, reflecting the global trend mentioned above. In addition, an increase in fishmeal factories on Norwegian trawlers has led to increased utilisation of unwanted catches but, as above, does not contribute to data about individual species.

Illegal catches

Illegal catches consist of those fish caught that the vessel had no legal right to take (i.e. due to being in closed areas or various gear regulations) or catches intentionally misreported upon landing (Pitcher et al. 2002). Intentional misreporting involves altering catch weights on official records, concealing illegal catches underneath legal catches in boxes, or exploiting difficulties in species identification. This is done to avoid prosecution for illegal fishing, catches being counted towards quotas, or get a better price than if it were legally landed. Fishing in illegal areas or periods requires a presence at sea to detect infringements, whilst intentional misreporting of landings requires portside inspections. Illegal catches are further complicated if one species is misreported as a different species, which results in a combination of under- and over-reporting. On-board fishmeal production or offal processing facilities can also be used to intentionally hide illegal catches. Methods for identifying the species composition of highly processed products require genetic techniques which are rapidly developing, but the detection of low-represented species is still particularly difficult and costly, rendering it currently unfeasible to routinely screen landed fishmeal (Vlachavas et al. 2019).

A study by Pitcher et al. (2009) found that there is poor compliance in fisheries globally. Across all countries, there are difficulties in controlling illegal fishing due to a mixture of poor policy implementation and lack of surveillance. The study assessed compliance with the United Nations Code of Conduct for Responsible Fisheries, finding that Norway had the highest score globally. Since 1990 when a new catch-monitoring system came into force in Norway, it has become increasingly difficult to misreport fish upon landing, especially for offshore fisheries (Gezelius 2006). The new system requires that daily catch logbooks and remote vessel monitoring at sea must match the information in sales notes completed when landing catches, reducing the risk of catches being misreported whilst at sea. Additionally, it is the joint responsibility between buyer and seller to report landings using approved weighing equipment. Finally, unannounced inspections mean that opportunities or incentives to misreport landings have been reduced, improving the reliability that official records accurately reflect what is landed (Gezelius 2006).

Discards

Discarding is caused by a complex combination of regulatory, environmental, and economic factors (Rochet and Trenkel 2005; Feekings et al. 2012; Pennino et al. 2017), all of which vary between fisheries and species. We therefore discuss the specific discard risks for different species groups in the next section. Discarding is further characterised by the conscious decision of skipper or crew to discard. Although discards can be reduced through regulations, improvements in gear selectivity, and improved utilisation of catches, some unwanted bycatches remain unavoidable. Fishing gears are seldom perfectly selective, and there is always the risk of non-compliance. In most cases, a discard ban will reduce discarding compared with fisheries without any discard regulations (Karp et al. 2019), but in worst-case scenarios a ban could increase the risk of discarding if monitoring and control is insufficient (Borges et al. 2016) or if additional management methods do not address any new problems that a discard ban creates (Pennino et al. 2017).

Slipping is considered as a type of discarding in this review because, like general discarding, it occurs during the hauling process, involves a decision by the skipper, and can result in high mortality rates (ICES 2020). Slipping most often occurs in purse seine fisheries as fishing strategies are more targeted towards very specific species and size groups, and catches are larger such that only a small number of hauls are needed to reach quota limits. As catches can be sampled before hauling the entire net, slipping becomes a solution to avoid undesirable catches. Slipping also occurs in trawl fisheries, but this is most commonly due to safety concerns, such as excessively large catches, damaged gear, or poor weather conditions. However, these issues are easier to mitigate as technology has developed.

Species-specific considerations

Different species groups are at risk of misreporting for different reasons and have different degrees of conservation concern (Hall 1996). Estimation procedures and output requirements will differ depending on the species and the need for estimating unreported catches (Anon 2003; Punt et al. 2006; Stock et al. 2018). It is therefore necessary to explore how catches can be categorised and what risks they are exposed to in order to determine the appropriate estimation procedure.

Target species

Due to their commercial value, target species typically undergo stock assessments to regulate their harvesting to achieve long-term sustainability. Therefore, one of the main goals for estimating unreported catches is to improve the accuracy of catch data used in stock assessments. Perretti et al. (2020) suggested that unreported catches should be accounted for, even if there is only a small possibility of their occurrence. This is based on evidence that the largest biases occurred when unreported catches were ignored, compared to accounting for them when they were not present. Rudd and Branch (2017) found that constant misreporting of catches can still produce sustainable estimates of recommended catches, but if misreporting varies over time, then estimates of important parameters become more inaccurate, and catch recommendations become more sensitive to the reporting rate. As a result of poor information on unreported catches, stock assessments can assume a constant value based on expert knowledge or long-term averages. However, this can introduce unknown biases in many aspects of a stock assessment. Whilst it is important to account for unreported catches, a constant rate will hide temporal trends, and may be unwillingly detrimental to the stock assessment.

Target species are generally included under a discard ban as they are typically subject to quota regulations. As a result, they are particularly vulnerable to high-grading, where lower value catches are discarded to make space for those with higher value to maximise the value of quota (Kelleher 2005; Batsleer et al. 2015). The risk of high-grading increases when approaching the quota limit, as a fisher aims for the highest return on the remaining quota. It can also be influenced by seasonal restrictions, minimum size requirements, low market value and storage restrictions during a trip (Batsleer et al. 2015). Despite the complex drivers behind high-grading, it results in the discarded portion having a different size distribution to the portion landed (Batsleer et al. 2015). Whilst high-grading is often based on the minimum landing size (Batsleer et al. 2015), it can also result in discarding of sizable fish if a vessel is actively targeting the largest of individuals (Stratoudakis et al. 1998). This was the case in Norwegian Barents Sea fisheries prior to the discard ban, where high-grading was legal.

Once the target species quota is filled, discarding will not be size selective as all catches of that species must be discarded to avoid penalties (Batsleer et al. 2015). This is especially relevant to 'choke' species, a species with low quota that when reached can force a vessel to stop fishing early, even though quotas for other species are available. Over-quota discarding involves large amounts of fish being discarded occasionally, as they are dependent on remaining quota, catch composition and available space on board. Aside from regulatory discarding behaviours listed above, a target species would otherwise be discarded only if damaged. This can occur from the prolonged soaking of passive gears leading to decay or predation, or the overcrowding in the codend of a trawl. Depending on the gear type, species and environmental conditions, damages may or may not be size based (Veldhuizen et al. 2018).

It is particularly in age- or length-based stock assessments where high-grading needs to be considered. Whether assuming a flat rate of discarding across all size groups, or constant size-based discarding across years, not accounting for the high variability in discarding of smaller size groups between years can mask annual variations in recruitment (Anon 2003; Dickey-Collas et al. 2007; Cook 2019), restricting the ability to detect strong incoming year classes that do not appear in reported landings (Punt et al. 2006). However, Punt et al. (2006) showed that if it is over-quota discarding that is the main cause of discarding, then it is unnecessary to account for size-based discarding patterns in the model. Instead, discards have the same length composition as landings so they can be combined to provide total catch estimates. Where both drivers are acting simultaneously, Cook (2019) demonstrated that only accounting for size-based discarding is inadequate if over-quota discarding is also occurring, which can account for as much as 40% of catches.

Justifying the assumption of either negligible or constantly unreported catches is especially important in multinational fisheries in Europe where each country contributes catch data to stock assessments. The magnitude of biases introduced by such assumptions depend on the relative contribution to total catches by that nation. Species with migratory behaviour may be vulnerable to different national fisheries at each life stage. As a result, the need to account for unreported catches of smaller fish (Anon 2003) would become the responsibility of nations whose fisheries overlap with nursery grounds, where the risk of high-grading is higher.

Bycatch species

Discarding of bycatch species with commercial value is primarily driven by market prices and storage space during trips but they can also be vulnerable to high-grading if subject to quotas (Batsleer et al. 2015), as well as becoming choke species if that quota is low relative to other species caught. There is also the risk that non-quota species are used to misreport species with limited quota. Commercial species that do not undergo detailed stock assessments may still be managed for their long-term sustainability. In these cases, size-based estimates may not be necessary, but total catches or numbers landed are still required to quantify total fishing mortality.

Non-commercial bycatches, sometimes referred to as 'incidental' catches, are those species that fishers have no intention of catching. Fish in this group can either be directed to fishmeal or discarded, creating a high risk of being unreported. Some of these species could have potential commercial value but are discarded or landed as fishmeal because there is currently no market for them. In these situations, quantifying unreported catches would help to assess the potential to develop a targeted fishery. New knowledge on catches could compliment scientific survey data to build a stock assessment which would provide evidence for a sustainable fishery. This would increase the value of the product, improve utilisation, and may help relieve pressure on more heavily fished alternatives if developed sustainably. Incidental catches also include endangered, threatened and protected species such as marine mammals, seabirds and sharks, and 'charismatic' species (Hall 1996) which when caught as bycatch can create a negative perception of the fishery (Gray and Kennelly 2018) and be a strong factor in influencing discard policy (Bellido et al. 2011).

Inaccurate estimates of unreported catches of non-commercial bycatch species will impact on management decisions, sustainability certifications for fisheries, and national import requirements. Management of unwanted catches is focussed on their avoidance under the Norwegian discard ban, so an estimation should consider the factors that influence their capture. For example, Cosandey-Godin et al. (2014) identified that bycatches of Greenland shark (*Somniosus microcephalus*) were confined to small geographical areas for the duration of each fishing season, but that these areas shifted between years, indicating that active spatial management is necessary to reduce bycatches. Sex- and age-biases are common in estimations of seabird bycatch (Gianuca et al. 2017), as they may influence their habitat or feeding behaviour, which in turn could affect their vulnerability to fishing gear. When monitoring the bycatches of non-commercial species to assess biodiversity and ecosystem function, neglecting fisheries bycatches will lead to an over-optimistic view of sustainability.

A fishery-based estimation of unreported catches

Based on various expert workshops and national reporting systems, it is commonly agreed that it is best to estimate unreported bycatches and discards by fishery (FAO 2015; ICES 2007a; NMFS 2011; Kennelly 2020). Framing the issue of unreported catches in a fisheries context allows for the consideration of unique dynamics and the broader ecosystem. For example, the management actions to reduce discards on one species may have a negative effect on mortality of other species through displacement (Gilman et al. 2019). A fishery-based approach will also complement the structure of sustainability certification assessment. Nevertheless, catch data requirements can differ between stocks depending on the selected assessment model and data availability. Therefore, for estimates of unreported catches to be useful, they should be of a similar type as those used in the stock assessments (Anon 2003), or appropriate for the available management options. This means that whilst estimations should be fishery-based, they should not disregard potential variations between species which would influence data collection requirements and the estimation procedure.

The management framework developed in Norway since the discard ban (Gullestad et al. 2017) provides the foundations for a fishery-based estimation of unreported catches. Fisheries are continuously assessed to prioritise issues such as the gear selectivity of different species groups and direct consideration of discards. Individual stocks also receive a similar assessment, which help to identify individual risks and demand for further knowledge for specific species. Norwegian stocks are also classified based on their economic importance and management objectives (Table 1). Within the table it is important to note that some species of low economic importance are grouped together due to limited knowledge. Estimates of unreported catches of individual species within these groups could help to distinguish them as a defined stock for targeted management.

Difficulties in enforcement and surveillance at sea mean that there is still a continued risk of discarding under the Norwegian discard ban. As a result, it is likely that discarding is still the main source of unreported catches in many fisheries. Improvements in the Norwegian reporting system and at-sea surveillance by the Norwegian Coast Guard and Directorate of Fisheries in recent decades have reduced the risk of discarding, illegal catches, and misreporting (Gezelius 2006; Gullestad et al. 2015). In 2019 the Norwegian Coast Guard conducted 1138 inspections and 738 aircraft surveillance hours with long range photo and video recording (Anon 2020). The use of drones and aircraft surveillance has greatly increased the ability to observe fishing vessels without detection. Nevertheless, there is always some degree of risk of illegal fishing. We have also argued why low-resolution reporting of fishmeal and certain species groups (e.g. sharks and rays) should be classified as unreported catches, even though they have been reported. Therefore, where there are no direct observations of discarding, caution should be used when interpreting the sources of unreported catches.

In fisheries using on-board fishmeal production, it is misleading to assume that unreported catches are a result of discards. Fishmeal production is a positive alternative to discarding, but can still be a source of unreported catches, so acknowledging the contributions will help to improve reporting requirements. Even with direct observations of discarding, it may be important to quantify the mortality of discarded fish, considering the exemption for discarding of live fish under the Norwegian discard ban. Discard survivability can be considerably higher in coastal fisheries where handling times are shorter (ICES 2020), whilst survivability of slipped catches in purse seine fisheries is highly variable, depending on a much wider range of factors, related both to fishing practices and environmental parameters (Tenningen et al. 2012, 2019; Gilman et al. 2013; ICES 2020). In such cases, contributions of discards to total fishing mortality may be overestimated if 100% mortality is assumed. In both these examples, poorly informed interpretations of results could be detrimental to the public image of the fishery and could lead to misguided management and enforcement decisions.

Data collection

The various methods for collecting data on bycatches and discards have been discussed extensively (ICES 2000; Cotter and Pilling 2007; Faunce 2011; Suuronen and Gilman 2020), providing a consesus on many of the benefits and limitations. However, more recent discussions on fisheries data collection under a discard ban (e.g. Kraan et al. 2013; Mangi et al. 2013; James et al. 2019) encourage a new evaluation of methods to address the influences of a ban and the consideration of novel methods and technologies. In this section, we gather the available data sources in Norwegian fisheries, as well as addressing data collection methods not currently used in Norway. Considering the limitations of the discard ban, we evaluate their ability to provide reliable data for estimating unreported catches, taking into account practical and social considerations.

Scientific observers

By far the most trusted method of sampling catches globally is by using on-board scientific observers (Anon 2003; Kelleher 2005; ICES 2007a; Suuronen and Gilman 2020). They are the major source of fisheries data collection in many countries (Karp et al. 2019), such as in the USA where numerous fisheries have achieved 100% coverage (NMFS 2011). Their benefits include the ability to gather a broad range of

Table 1 Summary of Norwegian stock classifications. Adapted from Gullestad et al. (2017)

Category	Type of stock	Contribution to total Norwegian first-hand value (%)	Management objectives
1	Economically most important marine fish stocks	90	Economically optimal long-term sustainable yield
2	Stocks of some economic importance, but about which information is scarce	5–7	High and, if possible, stable long-term sustainable yield
3	Stocks of low economic importance and non-commercial species	3–5	Ensure biodiversity and ecosystem function
4	Alien species	0	Reduce stock
0		0	Unsettled

data including catch composition, biological sampling, post-release survival and species identification (Suuronen and Gilman 2020), all of which can be collected based on a well-defined statistical sampling design to allow for a simple estimation procedure (Lohr 2010). Notwithstanding the above, the presence of an observer may influence fishing behaviour, known as the observer effect (Benoît and Allard 2009), whilst rejections or vessels being unsafe for observers could potentially bias the representativeness of sampled vessels. These effects are likely to be increased under a discard ban, where the presence of an observer would increase the risk of changing behaviour if the observer could witness illegal activity.

Many observer programmes worldwide require observers to report illegal activity on-board (Ewell et al. 2020). Arguments for merging scientific and monitoring roles include the moral obligation to report illegal activity, and improvements in compliance (especially with 100% coverage). However, for programmes focussing on unreported catches under a discard ban, there is an argument for the separation of roles (Cotter and Pilling 2007; Mangi et al. 2013). Even where observations are purely scientific, there could still be concerns from fishers about the later use of such data that could influence fishing behaviour or data quality. A review of 17 mandatory scientific observer programmes worldwide by Ewell et al. (2020) found that all programmes have issues with some aspect of the safety of their observers, regardless of the responsibility to monitor compliance. This includes a lack of measures to address intimidation, obstruction, and blackmail, but at worst, to investigate the disappearance or death of observers at sea. The risks to observer safety and welfare will be mitigated if observer roles are separated, but it is nevertheless important to consider that the presence of the discard ban will likely have negative effects on data quality from such programmes.

Higher observer coverage can reduce bias in estimates of unreported catches, but increasing the coverage without addressing rejection rates may weaken this improvement, or at worse increase bias (Lohr 2010). Increasing coverage is restricted by the high costs involved in maintaining an observer programme (Borges et al. 2004; Mangi et al. 2013). This is particularly the case in Norway where implementing an extensive scientific observer programme has been previously seen as logistically difficult, particularly for smaller demersal vessels. The extensive coastline has many landing sites that are separated by long fjords and mountains, making harbour access difficult for observers.

Remote electronic monitoring

The use of remote electronic monitoring (REM) is rapidly developing as an alternative to at-sea observers. For example, most recently REM programmes have been developed in commercial fisheries in Australia to improve the reliability of data from industry logbooks whilst reducing costs (Emery et al. 2019). Improved data reliability is also the reason for numerous European countries trialling REM in response to the landing obligation (Needle et al. 2014; Ulrich et al. 2015; James et al. 2019). Despite the infancy of REM technology, it is broadly seen as a vital tool in the future of fisheries monitoring (van Helmond et al. 2020), with its efficacy demonstrated as a mandatory requirement (Emery et al. 2019).

Nevertheless, James et al. (2019) highlighted that REM cannot provide physical samples such as otoliths for age determination, or data on maturity and sex, all of which can be necessary for stock assessments. Therefore, any data collection programme that uses REM must also include at least some form of human sampling.

Except for a vessel monitoring system, Norway does not have an REM programme for either the scientific monitoring, control or enforcement of catches. Part of the reason is due to technological limitations and high costs (NOU 2019), although both will likely improve as the technology develops (Suuronen and Gilman 2020). However, a more fundamental reason for a lack of uptake surrounds privacy concerns (NOU 2019), which is a serious barrier in the acceptance of REM programmes.

Enforcement and surveillance sampling

The Norwegian Directorate of Fisheries runs the Monitoring and Surveillance Service (MSS), an on-board observer programme for control purposes, which is divided into two categories. Observers can observe passively, gathering data on gross catches whilst the vessel is undergoing normal fishing activity, or they can hire a vessel for a specific objective, such as to identify bycatch hotspots for real-time closures. When MSS observers are passively observing, the observer effect could increase as skippers are concerned about reasons for the data collection. When vessels are hired, data do not represent normal fishing as samples will be clustered, confined to certain areas and times, and possibly contain more bycatches. However, if observations overlap with the active fishery, their representativeness could be justified.

The Norwegian Coast Guard also gathers data on catch compositions through at-sea enforcement inspections. Inspectors board vessels during the hauling procedure so that the skipper has selected the fishing ground without prior influence of the inspection, but vessel selection may be biased by a risk-based enforcement strategy. Alongside comparing logbooks to catches on board, inspectors take a representative sample of length measurements for commercial species to determine if the current haul contains a high proportion of undersized fishes.

MSS and Coast Guard inspectors are obliged to report any illegal activity they observe, making it highly unlikely for discarding to occur in their presence. Nevertheless, MSS and Coast Guard sampling is done on gross catches so still offer relevant information for estimating unreported catches through comparison with reported catches from vessels in the same area and time. An estimation of total retained catches in the Norwegian Economic Zone by Aanes et al. (2011) used Coast Guard inspections, stating that vessel selection is based solely upon the proximity to the pre-defined patrol route. Passive sampling by the MSS was used as the primary data source for the prediction of historical cod bycatch in the Barents Sea shrimp fishery (Breivik et al. 2017). Potential observer effects were deemed to be negligible due to the nature of the monitoring programme, but they did highlight that such assumptions should be reconsidered if the method is transferred to other fisheries.

Self-sampling

An alternative to observer sampling is self-sampling of catches by fishers, either throughout the entire fleet or by a defined group of vessels, known as a reference fleet (or study fleet). Mangi et al. (2013) distinguishes a reference fleet from other forms of fisher self-sampling by its enhanced data collection role. The Norwegian Reference Fleet is a collaboration between the Institute of Marine Research (IMR) and fishing industry, in which active fishing vessels are paid to collect data about their fishing activity and catches during normal fishing operations. It is divided into a coastal and offshore segment, covering both demersal and pelagic fisheries using gears such as trawls, purse seine, Danish seine, gillnets, longlines and traps.

The Norwegian Reference Fleet offers a direct source of information about discards as they are explicitly reported in samples. Coastal vessels began recording discards in 2005, whilst offshore vessels began in 2019. Prior to 2019, offshore vessels recorded gross catches. Sampling protocols differ between offshore and coastal vessels, and between gears, but the general routine involves constant reporting of landed catches and fishing activity, with biological sampling and reporting of discards (or gross catches) at regular intervals (Clegg and Williams 2020). Purse seine vessels also report details of slipping events.

All data recorded by the Norwegian Reference Fleet are property of IMR and are physically isolated from other catch records. An agreement between enforcement and surveillance authorities, IMR and fishers ensures that data shall not be requested for inspection or enforcement purposes. Even though this agreement is not legally binding, there have been no incidences where the agreement was compromised in the history of the programme, creating a trustful environment for fishers. This trust is core to the effectiveness of the programme. Reflecting upon the history of self-sampling programmes in New Zealand (Starr 2010), USA (Johnson and van Densen 2007), Ireland (Hoare et al. 2011; Lordan et al. 2011), the United Kingdom (Mangi et al. 2018) and the Netherlands (Kraan et al. 2013), long-term success relies on maintaining commitment and a strong communication channel between fishers and scientists. With membership in a reference fleet comes ownership in the scientific process, improving two-way support and communication between scientists and fishers and promoting transparency, which in turn will benefit other stakeholders, such as fisheries managers.

To maintain high quality data in the Norwegian Reference Fleet, IMR offers regular training, and IMR staff are assigned to vessels to maintain the sampling programme, regularly visiting vessels and checking incoming data. These data undergo the same quality assurance procedures as scientific survey data before being added to the database. One risk to data quality in long term self-sampling programmes is sampling fatigue (Hoare et al. 2011; Mangi et al. 2018). To alleviate this, the Norwegian Reference Fleet offers four-year contracts to vessels with direct monetary payment for sampling in compensation. An external evaluation of the Norwegian Reference Fleet by Bowering et al. (2011) concluded that based upon these quality assurance procedures, the programme meets the fundamental needs for effective scientific sampling of catches.

The reliability of self-sampling data has been open to question more than data collected by independent observers (Mangi et al. 2013). Based on scientific principles, data collectors should be disinterested in the scientific process. We must therefore acknowledge that fishers collecting the data may have a conflict of interest in the results from the data. Without regular quality control and validation, there is no direct evidence that proper, unbiased sampling protocols are consistently followed. Kraan et al. (2013) concluded that acceptance of self-sampling data by scientists can be hindered by a lack of trust in how the data are collected. The best practice for statistical data validation is to compare self-sampling data with a secondary source of data of known reliability (Fox and Starr 1996; ICES 2007b; Faunce 2011; Kraan et al. 2013), such as from scientific observers, remote electronic monitoring or scientific surveys. Importantly, such validation needs to be considered at all temporal scales to ensure that data quality is consistently maintained (Lordan et al. 2011) such that users have confidence in the data (Bell et al. 2017).

Whilst the Norwegian Reference Fleet maintains a strong quality control system, little has been done to validate it and there is no routine procedure in place for comparison with other reliable data sources. There is potential to investigate if data quality changes when IMR staff are on-board. Similarly, inspections by the Norwegian Coast Guard or passive observations by the MSS are done by independent observers and could therefore offer a suitable comparison. Nevertheless, qualitative evidence of reliability is available through multiple studies estimating the bycatch of species of high conservation importance, namely seabirds (Fangel et al. 2015; Bærum et al. 2019) and porpoises (Bjørge et al. 2013) in coastal gillnet fisheries. Reporting of seabirds and sea mammals by the Norwegian Reference Fleet is notably higher than through official reporting channels, indicating a greater willingness to record sensitive data for scientific purposes.

A fundamental aspect of a reference fleet is its representativeness of the wider fishing fleet (Mangi et al. 2013). The vessel selection process in the Norwegian Reference Fleet limits the use of a truly random sampling design, as it is legally required to follow a publicly transparent tender process (Clegg and Williams 2020). Vessels can voluntarily submit applications, which could introduce bias in vessel selection. Willingness to participate will increase the reliability of data but, as is the case with rejections in observer programmes, vessels willing to participate in a reference fleet may behave differently to those unwilling. To account for this, contracts are awarded based on gear and vessel specifications, fishing patterns and coverage to mitigate bias and ensure stratification throughout fisheries. For a non-random vessel selection where the statistical properties of the sample are unknown, using statistical tests to assess representativeness is not recommended (Anon 2003). Instead, general comparisons in vessel characteristics and fishing behaviour of sampled vessels can be compared to the wider fishery to determine representativeness on a case-by-case basis (Anon 2003). Such studies have been done for the Norwegian Reference Fleet in general (Bowering et al. 2011), but individual studies should be done prior to implementing programmes in specific fisheries. For example, a comparison of estimates of seabird bycatches in the Norwegian coastal gillnet fishery using Norwegian Reference Fleet data and access-point surveys of the broader fleet (Fangel et al. 2015) yielded identical results, giving evidence for the representativeness of reference fleet data for the reporting of non-commercial and controversial bycatches.

Industry data and mandatory reporting

Under a discard ban, official landings statistics are a record of all species landed by commercial vessels and are therefore the reference to which unreported catches are compared. Norwegian vessels must fill out a daily logbook which records information about individual hauls, including locations and total weights of catches per species. Upon returning to port, a landing note is generated which contains all catches on that trip. Through the daily logbook and landing notes, catch and effort data are available for the entire Norwegian fishing fleet.

Regarding data on gross catches, the data collection methods discussed so far have focussed on active sampling programmes which require some form of human observation. However, modern fishing vessels use various electronic instruments to routinely gather data whilst fishing, either for commercial purposes or for mandatory reporting. The most well-known example involves satellite tracking of vessel movement, which is now widely used for control, surveillance and for scientific research (e.g. Aanes et al. 2011). Other sources of industry data include weighing of catches in the codend or on platform scales, and onshore grading machines used in fish markets to grade catches before sale (Mangi et al. 2013).

There are continual difficulties in biological sampling of catches in Norway, leading to large uncertainties in age and length compositions of catches for many fisheries (Bowering et al. 2011). An intercept sampling programme ran by IMR samples landings at specific harbours north of 62 °N latitude, although the programme focusses mainly on coastal vessels landing whole fresh fish. For vessels with on-board factories landing processed and frozen catches, intercept sampling requires the defrosting of products which affects their value, making it unfeasible. Instead, there is the potential to obtain size-based data of fishes during the grading process on board factory vessels before they are frozen, when species are identified then sorted into weight grades. Importantly, the weights of individual fish are recorded for each haul, offering a higher resolution of information necessary for accurate size distributions both spatially and temporally (Plet-Hansen et al. 2020).

There are aims to develop technology to monitor the entire harvesting process in Norway (NOU 2019). This involves automatic recording of catches at the earliest possible stage after hauling, including species identification and individual weights. Such a system would vastly improve knowledge on total extractions from fisheries and reduce the need for estimation studies if there is evidence for high compliance and reliability of data. However, until this goal is met, data from the on-board grading process could provide size-based information on landed catches which can be compared with gross catches to infer unreported catches.

Scientific surveys

Where fisheries-dependent data are unavailable or are inadequate due to reasons such as rare encounters or poor coverage, scientific survey data are a possible alternative (Fox and Starr 1996; Cook 2013). If a survey overlaps with the target fishery in both space and time then it could offer systematic, random sampling robust enough for statistical analysis (Fox and Starr 1996), albeit with caveats. Scientific surveys are very expensive compared to fisheries-dependent data, restricting their spatial and temporal coverage. The survey fishing gears commonly use finer meshed nets to catch a broad range of size classes and species, and towing times are often shorter. If these factors can be accounted for, then scientific survey data can be used in place of, or to enhance, fisheries-dependent data.

Opportunities can arise where specific survey gear has been calibrated against commercial gear in the fishery, allowing for appropriate conversions (e.g. Mayo et al. 1981; Hylen and Jacobsen 1987; McBride and Fotland 1996; Dingsør 2001b). However, routine estimations would require regular calibration studies to reflect developments in gear technology and fishing patterns by the commercial fleet. Otherwise, conversions can be based on theory (Heath and Cook 2015), or under the strong assumption of 'knife-edge' size selection of species at a certain length such as the minimum landing size (Mayo et al. 1981), which will introduce further uncertainty. Scientific survey design is generally of a high quality relative to fisheries-dependent sampling programmes, as scientific surveys can be highly controlled, and involve less risk and opportunism. However, the calibration methods required due to the use of non-commercial gears outweighs these benefits. Updating calibrations is not sustainable in the long-term for regular estimates of unreported catches, especially as modern fishing technology rapidly develops. Therefore, studies that have used this approach have acknowledged it is only useful in the absence of direct observations of fishing activity (McBride and Fotland 1996).

More recently, unreported catches have been estimated directly in the stock assessment modelling process, using scientific survey indices and reported catches (Hammond and Trenkel 2005; Bousquet et al. 2010; Heath and Cook 2015; Cadigan 2016), and can also incorporate observations of discarding if available (Cook 2019). In extreme cases where catch reporting is deemed highly unreliable, it can be disregarded completely in favour of an assessment using only research survey data (Cook 2013). Incorporating estimations into the stock assessment model bypasses the need to calibrate fishing gears and will benefit from continual developments in modelling tools and techniques. Whilst improvements could be made to how unreported catches are incorporated into stock assessment models, Cook (2013) acknowledges such a method should not be seen as a replacement for methods incorporating catch data, but instead be an additional tool for comparison where catch data are unreliable.

Utilising multiple data sources

Direct observations still provide the best opportunities for estimating unreported catches, despite the difficulties in observing normal fishing activity at sea under a discard ban. Self-sampling of catches by the Norwegian Reference Fleet alleviates the issue of trust, as data shall not be used for enforcement purposes, and has improved the relationship between science and industry such that results are accepted. Control and enforcement data should not be completely disregarded as a viable data source, despite issues of vessel selection and observation biases. They can serve to enhance scientific sampling programmes where data gaps are present and help particularly in closed areas when identifying bycatch hotspots. The appropriateness of surveillance or enforcement observations need to be determined for each study, requiring expert knowledge of the sampling methodologies to justify their use. Finally, scientific survey data are beneficial only where direct information is unavailable or unreliable (Cook 2013; Heath and Cook 2015), although there are examples of benefits where direct observations of discards have been included in the stock assessment model, utilising both fisheries-dependent and -independent data sources (Punt et al. 2006; Cook 2019).

New data collection methods should also be considered to improve data quality, either as an improvement to current sampling programmes (e.g. REM technologies) or where data are not available. For example, on offshore pelagic vessels, enclosed catch systems limit the opportunities to sample catches at sea. To gain sufficient information in this situation, catch volumes could be monitored using sensors to monitor the pipe system and storage tanks, with complimentary portside sampling providing information on catch composition.

Estimation procedure

A good estimation of unreported catches should be unbiased, precise, and simple (ICES 2007a). However, the scope and design of a study will affect the extent to which this goal can be met. A well-chosen estimator can account for various sources of bias and provide an accurate estimate of the uncertainty. Conversely, a poor estimator can introduce further biases and give a misleading view of uncertainty. In this section, we consider how all the themes discussed so far can influence the choice of the best available estimator.

Design- and model-based approaches

Estimates of unreported catches or discards can be obtained using standard formulae for extrapolations based on defined sampling programmes (e.g. Cochran 1977; Lohr 2010), known as the design-based approach. Design-based estimators rely on probabilistic sampling to ensure that the sample is representative of the population (Lohr 2010), but it is realised that high rejection rates or vessels being unsafe for observers mean that the samples can drift away from a truly probabilistic selection (Table 2). Alternatively, estimates of unreported catches or discards can be obtained using a modelling approach by estimating a set of unknown parameters that explain variations. Model-based estimators do not require probabilistic sampling, but can benefit from randomisation of important covariates, although it is necessary for the range of each covariate to be adequately covered in samples (Cotter and Pilling 2007). Where there are direct observations of discards, then these samples can be extrapolated using either a design- or model-based approach. In the absence of direct observations, then gross catches can be extrapolated to get an estimate of total catches in the fishery, then compared to reported catches to infer misreporting.

General applications of design-based estimators have been adapted for estimating discards and bycatches, producing best practice guidelines for various types of sampling (e.g. Anon 2003; ICES 2007a; Vigneau 2006). They acknowledge that the optimal procedure is highly case-specific, meaning there cannot be a simple, straight-forward method applicable generally. It is therefore necessary for every new study to identify the suitable estimators based on the sampling design and assumptions, then systematically compare them (ICES 2007a). It is common to assume that discards are proportional to an auxiliary variable such as catch or effort, allowing for extrapolation using a ratio estimator (Cochran 1977). However, a review by Rochet and Trenkel (2005) found that in all 17 case studies they considered, both catches and effort were either not influential or had a non-linear relationship with discards. In reality, studies are often constrained by data availability. The auxiliary variable required for extrapolation needs not only to be recorded during sampling, but also documented reliably for the entire fishing fleet. It is therefore possible that studies may only be able to use one procedure to obtain an estimate. In these cases, preliminary studies are still necessary to identify issues beforehand (Borges et al. 2005), as basing estimates on assumptions can introduce unknown bias and uncertainty.

Earlier workshops developing estimation methodologies did not give a large consideration to model-based estimators, mainly due to the absence of suitable case studies (ICES 2000, 2007a). However, over the last two decades there have been advances in techniques for dealing with complexities such as clustered sampling (Harrison et al. 2018), low encounter rates (Martin et al. 2005), spatial–temporal correlation (Rue and Martino 2009) and their extensions to multispecies estimations (Thorson et al. 2017). The appropriate application of these methods can result in reduced bias (Breivik et al. 2017) or improved precision (Stock et al. 2018). These methods have also seen improved computation times and more open-source support, making them more accessible to fisheries studies.

Factors affecting the choice of estimator

If high-grading is to be investigated, then a size-based estimation is necessary. Liggins et al. (1997) compared mean lengths of retained fish sampled at sea and landed catches. Although this was to detect bias in sampling of retained catches at sea, applying the same analysis with gross catches at sea would provide a method for detecting high-grading. This was used by Pálsson (2003) to compare the size distributions of aggregated samples at sea and onshore to model the probability of discard at length (see also Borges et al. 2006), which can then be extrapolated to quantify unreported catches in the entire fishery. Alternatively, multiple fish lengths or ages can be modelled simultaneously using a multivariate modelling approach (Thorson 2019). The Norwegian Reference Fleet is currently the primary source of age- and length-based data in many Norwegian fisheries. An external evaluation of the programme (Bowering et al. 2011) collated comments from various stock assessment working groups to identify that low sampling coverage of vessels and for certain gear types has impacted on the precision of estimates. Where age-length keys are used to estimate catch at age from fisheries, this has resulted in difficulties in estimating catches for those size-groups that are under-represented. The port intercept sampling programme in northern Norway only covers coastal fisheries, and is merged with Norwegian Reference Fleet data to improve size-

Table 2 Summary of design- and model-based solutions to issues surrounding the estimation of unreported catches

Issue	Approach	Solution	Limitations	References
Non-random selection of vessels	Design-based	Assume random selection and apply the appropriate estimator	Bias can be introduced if assumption is not met	Cochran (1977) and Lohr (2010)
	Model-based	Include vessel characteristics as fixed effects (e.g. engine power, vessel length)	Vessel characteristics may not be available for sampled vessels	Batsleer et al. (2015)
		Include vessel as a random factor to account for the hierarchical nature of the data	Requires more than five groups and relatively balanced sample sizes across groups	Harrison et al. (2018)
Unsampled strata	Design-based	Impute values for missing strata based on similar strata	Risks of misusing results if it is not clear which strata were imputed	Lohr (2010)
			Poor assumptions of similarity may introduce bias	
		Ad hoc or objective collapsing of strata	Unknown biases in subjective approaches	Stratoudakis et al. (1999) and Anon (2003)
	Model-based	Include spatial variables in the model (e.g. depth, distance from coast) to be able to predict in unsampled strata	Requires knowledge of environmental drivers and that the relevant data are available	Bremner et al. (2009)
		Account for spatial correlation to 'borrow' information from other sampled strata	Requires advanced statistical knowledge	Rue and Martino (2009), Cosandey-Godin et al. (2014) and Breivik et al. (2017)
			Requires coordinates of samples	
Multiple species or fisheries comparison	Design-based	Assume that catches or discards are correlated with the same auxiliary variable across all species or fisheries	Unknown biases introduced for all cases where auxiliary variable does not have a strong linear correlation with unreported catches or discards	Kelleher (2005), Rochet and Trenkel (2005) and Pérez Roda et al. (2019)
		Apply multiple estimators to all cases to allow for more comparisons	Extreme differences may expose biased estimators but could still produce unknown biases	Kennelly (2020)
	Model-based	Assume the same explanatory variables influence catches or discards across all species	Poorer model fit from excluding potential drivers unique to individual species	Stock et al. (2018)
		Apply model selection procedures to select significant variables for each species	Unfeasible for a large number of species	Bremner et al. (2009)
Rare encounters	Design-based	Increase sample size	Cost is a limiting factor in the expansion of many sampling programmes	Borges et al. (2004) and Lohr (2010)
		Adapt sampling to account for rare events	Sampling programmes often aim to cover multiple species. Adapting the design may impact on the estimation of other species	Lohr (2010)
		Separate occurrences and non-occurrences using a delta-lognormal estimator	Misleading results if underlying distribution of non-zero occurrences is not lognormal	Pennington (1983)
	Model-based	Zero-inflated modelling techniques	Requires a firm understanding of the processes causing zero values	Martin et al. (2005)

Table 2 continued

Issue	Approach	Solution	Limitations	References
Size-based estimate	Design-based	Adapt design-based approach to extrapolate estimates by size	Difficulties in estimating uncertainty	Pálsson (2003)
	Model-based	Incorporate size-based variables into the model such as the probability of discarding at length	Dependent on the availability and reliability of such data	Borges et al. (2006)
		Model all length classes in a multivariate model	Multivariate modelling requires advanced statistical knowledge	Thorson et al. (2017)

based data for stock assessments. However, this is based on the assumption that all catches are landed, which requires an estimate of unreported catches to justify. Therefore, the quantification of high-grading is also restricted by the absence of size-based data on landings.

Multiple species estimations may be necessary in highly non-selective fisheries or when obtaining estimates for multiple fisheries for a national or global study. Comparisons can be made by using the same design-based estimator across all species or fisheries (Table 2). For example, global discard studies (Kelleher 2005; Pérez Roda et al. 2019; Gilman et al. 2020) assume a relationship between discards and reported landings, as landings data are more readily available than fishing effort. However, this relationship is not always justifiable (FAO 2015; Kennelly 2020), with discards being more often correlated with fishing effort. Therefore, in cases where both landings and effort are available, both should be used to allow for comparisons. For model-based estimators, a univariate approach can assume the same covariates are driving discarding across all species (Stock et al. 2018), but this is understandably not ideal for species with very dissimilar life histories or catch patterns. An alternative is to determine important drivers for each species (Bremner et al. 2009), which would improve accuracy, but could quickly become unfeasible as the number of species and covariates increased. Finally, multiple species can be modelled simultaneously in a joint species distribution modelling framework (Thorson et al. 2015, 2016). This addresses issues of multi-model approaches, whilst improving accuracy. The approach is particularly beneficial for rare or under-sampled species, where information on the co-occurrence of more frequently observed species can be used to improve accuracy of estimates.

Post-stratification is used due to the inability to select strata before sampling (as is true for the Norwegian Reference Fleet, and a likely scenario in many observer programmes), but it may result in certain strata being under-sampled. A model-based estimator allows unsampled strata to 'borrow' knowledge from similar strata where sample sizes are too small for a design-based estimate (Lohr 2010) (Table 2). Nevertheless, ad hoc solutions to poorly sampled strata are available for design-based estimators, such as collapsing the stratification, assuming values based on similar strata, or excluding the stratum from the study (Anon 2003). Stratification is partly based on the hypothesis that environmental conditions influence discards (Rochet and Trenkel 2005). Therefore, solutions to unsampled strata can cause misleading results and should always be justified (Stratoudakis et al. 1999). Any biases introduced from imputation would have little impact if strata were unsampled due to low fishing activity. However, if estimates for heavily fished strata must be imputed, then the imputation method requires a stronger justification.

Probabilistic sampling of rare encounters requires special adaptations in sampling design, which will likely not be accounted for in sampling programmes focused on the broader fishery (Table 2). This can either be in the form of sampling a rare population, such as an endangered species, or the observation of rare but extreme events (Lohr 2010), such as slipping of large catches in purse seine fisheries. Using standard formulae for common occurrences with rare encounters could result in biased estimates and an incorrect estimation of variance (Lohr 2010). Sampling can be adapted to account for this but could be impractical alongside the standard sampling programme for other species. Solutions include the delta-lognormal method (Pennington 1983), where zeros are treated separately to occurrences in the estimator, or zero inflated modelling methods (Martin et al. 2005).

The estimation of total mortality from slipping requires the consideration of more factors in addition to the estimation of rare events. The low number of total fishing operations in purse seine fisheries will alter assumptions about sampling coverage and representativeness compared to other fishing methods. For example, although Reference Fleets sample each vessel and fishing operation without replacement, low sampling coverage can allow for the assumption of replacement to allow for the use of simple estimators (Lohr 2010). However, this assumption may not hold in purse seine fisheries where there are relatively low numbers of vessels and fishing operations each year. Contributions to total mortality from slipping is highly dependent on a complimentary study on survivability. Depending on the timing of the slipping event, catch size and species, mortality rates can range from 1 to 100% (ICES 2020). It is difficult to accurately measure or estimate the weight of slipped catches before they are released (Tenningen et al. 2019). Therefore, a good understanding of mortality from slipped catches would first need to estimate the rate of slipping events, the total biomass of the slipped catches, and the survivability post-release. The diverse methodological and statistical requirements for estimating each of these steps may explain why slipped catches are understudied relative to other sources of unreported catches.

General issues of complexity should also be considered when communicating complex models to stakeholders. Poor communication can lead to misinterpretation, misuse, and mistrust of the results (Cartwright et al. 2016). When selecting a more complex approach, there is a responsibility to involve stakeholders during the modelling process. Scientists should also ensure that the decisions and assumptions are transparent and well-communicated, such that it does not restrict the ability for stakeholders to understand and criticise the results. There was previously an argument for considering the computation time of complex models. However, with advancements in computing power and software development, such run times are now measured in hours or minutes (Rue and Martino 2009; Cosandey-Godin et al. 2014; Breivik et al. 2017).

Performance of estimators

With advances in statistical modelling approaches, there is a strong case for using model-based approaches to estimate unreported catches. Another argument is the reduced dependence on the probabilistic sampling designs necessary for a design-based estimation (Cotter and Pilling 2007). The representativeness of probabilistic sampling may be compromised by rejections or inaccessible vessels, or the inability to do random sampling like the case of non-random vessel selection in the Norwegian Reference Fleet.

The benefits of design-based estimators are their versatility and simplicity, so for modelling to be justified, any improvements from increased complexity should outweigh the simplicity of a design-based approach (Stock et al. 2018). Despite the increasing popularity of modelling approaches, there is still no firm understanding of how they compare to simpler design-based methods. Both design- and model-based approaches can account for a wide range of complexities in an estimation (Table 2). In each case, there will likely be one approach that performs better, but this is dependent upon how such performance is defined.

A common measure of performance of an estimator is the trade-off between accuracy and precision (Amande et al. 2012; Stock et al. 2018). For commercial species, stock assessments require accurate estimates of total catches in the fishery, whilst the monitoring of catches of rare species over time favours precision over accuracy, as the relative changes are important in explaining their vulnerability to capture by fishing patterns over time (Stock et al. 2018). This has been demonstrated by Stock et al. (2018) and Breivik et al. (2017), who both compared spatial–temporal models to standard design-based estimators. Stock et al. (2018) found that model-based approaches performed best across the 15 species considered, despite a small increase in bias. Contrastingly, Breivik et al. (2017) found that a modelling approach reduced bias in estimates, but uncertainty was not estimated for the design-based estimators to allow for a comparison. Considering this trade-off can therefore be a useful tool for deciding the best estimator, taking into account also the factors discussed in the previous section and data availability.

Where unreported catches are estimated within a stock assessment model, there is not the same opportunity to gather multiple estimators for comparison. However, performance can still be evaluated through general best practices for model validation, such as through the reduction of total error in the model (Perretti et al. 2020), and the final model can be tested using well-established procedures such as simulation testing (Cadigan 2016; Cook 2019), cross validation (Heath and Cook 2015) and sensitivity analysis (Heath and Cook 2015).

Conclusions

This review has identified a range of best practices for estimating unreported catches which, whilst in the context of Norwegian fisheries under a discard ban, are framed to be relevant to other discard bans globally where similarities can be identified. We have explored a broad range of aspects related to the estimation of unreported catches, and therefore offer the main conclusions below:

(1) If there are no direct observations of discards, then unreported catches can be estimated by comparing gross catches with landings. This limits the interpretation of results and management recommendations for those studies which cannot determine the relative contributions of individual sources, or where survivability of discards should be considered.

(2) For estimates to be effective, their required use should be considered in the presentation of results. This includes considering the data structure in a stock assessment or current management plans, and good communication of accuracy and uncertainty.

(3) Unreported catches should be estimated on a fishery-by-fishery basis to effectively include fishery-related factors and account for potential consequences on management of other species.

(4) Self-sampling of gross catches and discards has the potential to address some of the data collection issues created by the discard ban. Cooperative research can improve trust and transparency between fishers and scientists, which in turn improve the acceptance of data and results (Johnson and van Densen 2007; Starr 2010; Lordan et al. 2011; Kraan et al. 2013; Mangi et al. 2018).

(5) Reliability of self-sampling is more open to question than for independent scientific observers. There are still concerns from the scientific community regarding the reliability of self-sampled data, which must be addressed statistically by comparing self-sampled data with another data source of known reliability.

(6) Studies can benefit from utilising multiple data sources, either to fill in data gaps or to increase observations, but potential biases should be considered.

(7) Representativeness of data should be assessed prior to each study to assess the risk of bias in estimates. Differences in regulations, harvesting strategies and sampling protocols make it unadvisable to generalise across fisheries.

(8) Model-based estimators should be applied, especially where non-random sampling designs have been applied. However, comparisons should be made with design-based estimators to justify the increase in complexity (Table 2). A useful method to determine the best estimator is the trade-off between bias and precision, which is in turn determined by the desired use of the estimate.

A fishery-based approach to estimating unreported catches can be readily incorporated into the Norwegian management system, which requires knowledge of total extractions of all species from fisheries, as well as graded objectives for individual fisheries, commercial stocks and bycatch species (Gullestad et al. 2017). Use of the fisheries and stock tables (Gullestad et al. 2017) should help to prioritise studies depending on their demand for estimates of unreported catches.

Various studies have estimated unreported catches in Norway for commercial species as both target species (Aanes et al. 2011) and bycatch (Breivik et al. 2017), as well as incidental catches of species with high conservation importance (Bjørge et al. 2013; Fangel et al. 2015; Bærum et al. 2019). They have utilised a wide variety of data sources and estimation procedures to extrapolate directly from sampled catches or infer from indirect sources. We argue that the Norwegian Reference Fleet has the greatest potential for estimating unreported catches in a wide range of fisheries in Norway. However, it will be necessary to consider multiple estimators to account for the various fleet segments, gear-specific sampling protocols and the characteristics of each fishery. Therefore, where methods are trialled then it should be considered where generalisations to similar fisheries are justifiable. Furthermore, methodologies should be reviewed at defined intervals to address changes in representativeness, sampling protocols, and advances in gear technology.

In considering the usefulness of Norwegian Reference Fleet data, the above recommendations for evaluating the representativeness of data need to be addressed. The vessel selection procedure in the Norwegian Reference Fleet aims for representativeness through expert judgement and random selection from eligible vessels. To assess the extent to which this process behaves like a simple random sample, a devoted study may help to explore the representativeness on a broader scale, whilst identifying those fisheries where the vessel selection procedure or sampling protocols could introduce bias.

The focus on self-sampling in this review is not without regard to the benefits of other methods, but rather due to the demand to identify and evaluate the data sources that are currently available in Norway. Following this, the benefits of REM (Emery et al. 2019) and industry data sources (Plet-Hansen et al. 2020) should be considered to improve future estimations. For example, incorporating REM into the Norwegian Reference Fleet would reduce workload to allow for more extensive sampling of hauls. Utilising data from fish grading systems on board factory vessels could address the current data gap in many Norwegian fisheries regarding detailed size distributions of landed catches (Bowering et al. 2011). The current mandatory reporting requirements generate size-based data which are too coarse for comparison with size distributions of gross catches from the Norwegian Reference Fleet.

Finally, the estimation of unreported catches from slipping is in a much earlier stage in Norwegian fisheries. This is partly because it involves multiple studies to understand the extent, scale, and survivability of slipping events. Sampling protocols in the Norwegian Reference Fleet include the recording of slipping events, but their suitability has not yet been determined. We therefore recommend investing in exploratory studies prior to a devoted estimation to address questions such as data requirements, appropriate sampling designs, and what approaches are suitable to synthesise the knowledge of scale and survivability to arrive at an estimation of total mortality.

Acknowledgements The authors thank Kotaro Ono, Hans Hagen Stockhausen, and two anonymous reviewers for their helpful comments on the manuscript.

Author's contribution Article originally devised and drafted by Thomas L Clegg. All other authors contributed to various versions of the manuscript.

Funding Open Access funding provided by Institute Of Marine Research. This review was funded equally by the Institute of Marine Research and the Directorate of Fisheries through taxation of first-hand sales of fish in Norway.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest to declare.

References

Aanes S, Nedreaas K, Ulvatn S (2011) Estimation of total retained catch based on frequency of fishing trips, inspections at sea, transhipment, and VMS data. ICES J Mar Sci 68:1598–1605. https://doi.org/10.1093/icesjms/fsr117

Amande MJ, Chassot E, Chavance P et al (2012) Precision in bycatch estimates: the case of tuna purse-seine fisheries in the Indian Ocean. ICES J Mar Sci 69:1501–1510. https://doi.org/10.1093/icesjms/fss106

Anon (2003) Workshop on discard sampling methodology and raising procedures. ICES Planning Group on Commercial Catch, Discards and Biological Sampling, Chalottenlund, Denmark 2–4 Sept 2003

Anon (2020) Forsvarets årsrapport 2019. https://www.forsvaret.no/aktuelt-og-presse/publikasjoner/forsvarets-arsrapport. Accessed 19 Dec 2020

Bærum KM, Anker-Nilssen T, Christensen-Dalsgaard S et al (2019) Spatial and temporal variations in seabird bycatch: incidental bycatch in the Norwegian coastal gillnet-fishery. PLoS ONE 14:e0212786. https://doi.org/10.1371/journal.pone.0212786

Batsleer J, Hamon KG, van Overzee HMJ et al (2015) High-grading and over-quota discarding in mixed fisheries. Rev Fish Biol Fish 25:715–736. https://doi.org/10.1007/s11160-015-9403-0

Bell RJ, Gervelis B, Chamberlain G, Hoey J (2017) Discard estimates from self-reported catch data: an example from the U.S. Northeast Shelf. North Am J Fish Manag 37:1130–1144. https://doi.org/10.1080/02755947.2017.1350219

Bellido JM, Santos MB, Pennino MG et al (2011) Fishery discards and bycatch: solutions for an ecosystem approach to fisheries management? Hydrobiologia 670:317–333. https://doi.org/10.1007/s10750-011-0721-5

Benoît HP, Allard J (2009) Can the data from at-sea observer surveys be used to make general inferences about catch composition and discards? Can J Fish Aquat Sci 66:2025–2039. https://doi.org/10.1139/f09-116

Bjørge A, Skern-Mauritzen M, Rossman MC (2013) Estimated bycatch of harbour porpoise (*Phocoena phocoena*) in two coastal gillnet fisheries in Norway, 2006–2008. Mitigation and implications for conservation. Biol Conserv 161:164–173. https://doi.org/10.1016/j.biocon.2013.03.009

Borges L, Zuur AF, Rogan E, Officer R (2004) Optimum sampling levels in discard sampling programs. Can J Fish Aquat Sci 61:1918–1928. https://doi.org/10.1139/f04-138

Borges L, Zuur AF, Rogan E, Officer R (2005) Choosing the best sampling unit and auxiliary variable for discards estimations. Fish Res 75:29–39. https://doi.org/10.1016/j.fishres.2005.05.002

Borges L, Zuur AF, Rogan E, Officer R (2006) Modelling discard ogives from Irish demersal fisheries. ICES J Mar Sci 63:1086–1095. https://doi.org/10.1016/j.icesjms.2006.03.017

Borges L, Cocas L, Nielsen KN (2016) Discard ban and balanced harvest: a contradiction? ICES J Mar Sci 73:1632–1639. https://doi.org/10.1093/icesjms/fsw065

Bousquet N, Cadigan N, Duchesne T, Rivest L-P (2010) Detecting and correcting underreported catches in fish stock assessment: trial of a new method. Can J Fish Aquat Sci 67:1247–1261. https://doi.org/10.1139/F10-051

Bowering R, Storr-Paulsen M, Tingley G et al (2011) Evaluation of the Norwegian Reference Fleet. Institute of Marine Research, Bergen. http://hdl.handle.net/11250/116704. Accessed 03 Jan 2020

Breivik ON, Storvik G, Nedreaas K (2017) Latent Gaussian models to predict historical bycatch in commercial fishery. Fish Res 185:62–72. https://doi.org/10.1016/j.fishres.2016.09.033

Bremner G, Johnstone P, Bateson T, Clarke P (2009) Unreported bycatch in the New Zealand West Coast South Island hoki fishery. Mar Policy 33:504–512. https://doi.org/10.1016/j.marpol.2008.11.006

Cadigan NG (2016) A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Can J Fish Aquat Sci 73:296–308. https://doi.org/10.1139/cjfas-2015-0047

Cartwright SJ, Bowgen KM, Collop C, Hyder K, Nabe-Nielsen J, Stafford R, Thorpe RB, Sibly RM (2016) Communicating complex ecological models to non-scientist end users. Ecol Model 338:51–59. https://doi.org/10.1016/j.ecolmodel.2016.07.012

Clegg T, Williams T (2020) Monitoring bycatches in Norwegian fisheries—species registered by the Norwegian Reference Fleet 2015–2018. Report 2020-8. Institute of Marine Research, Bergen, Norway

Cochran WG (1977) Sampling techniques, 3rd edn. Wiley, New York

Cook RM (2013) A fish stock assessment model using survey data when estimates of catch are unreliable. Fish Res 143:1–11. https://doi.org/10.1016/j.fishres.2013.01.003

Cook RM (2019) Inclusion of discards in stock assessment models. Fish Fish 20:1232–1245. https://doi.org/10.1111/faf.12408

Cosandey-Godin A, Krainski ET, Worm B, Flemming JM (2014) Applying Bayesian spatiotemporal models to fisheries bycatch in the Canadian Arctic. Can J Fish Aquat Sci 72:186–197. https://doi.org/10.1139/cjfas-2014-0159

Cotter AJR, Pilling GM (2007) Landings, logbooks and observer surveys: improving the protocols for sampling commercial fisheries. Fish Fish 8:123–152. https://doi.org/10.1111/j.1467-2679.2007.00241.x

Dickey-Collas M, Pastoors MA, van Keeken OA (2007) Precisely wrong or vaguely right: simulations of noisy discard data and trends in fishing effort being included in the stock assessment of North Sea plaice. ICES J Mar Sci 64:1641–1649. https://doi.org/10.1093/icesjms/fsm155

Dingsør GE (2001a) Norwegian un-mandated catches and effort. In: Zeller D, Watson R, Pauly D (eds) Fisheries Impacts on North Atlantic Ecosystems: Catch, Effort and National/Regional Data Sets. Volume 9, Number 3. The Fisheries Centre, University of British Columbia, Vancouver, pp 92–98

Dingsør GE (2001b) Estimation of discards in the commercial trawl fishery for Northeast Arctic cod (*Gadus morhua* L.) and some effects on assessment. Cand. Scient thesis, University of Bergen

Emery TJ, Noriega R, Williams AJ, Larcombe J (2019) Changes in logbook reporting by commercial fishers following the implementation of electronic monitoring in Australian Commonwealth fisheries. Mar Policy 104:135–145. https://doi.org/10.1016/j.marpol.2019.01.018

Ewell C, Hocevar J, Mitchell E et al (2020) An evaluation of Regional Fisheries Management Organization at-sea compliance monitoring and observer programs. Mar Policy 115:103842. https://doi.org/10.1016/j.marpol.2020.103842

Fangel K, Aas Ø, Vølstad JH et al (2015) Assessing incidental bycatch of seabirds in Norwegian coastal commercial fisheries: empirical and methodological lessons. Glob Ecol Conserv 4:127–136. https://doi.org/10.1016/j.gecco.2015.06.001

FAO (2015) Report of the expert workshop on the methodology to assess and quantify the extent and impact of fisheries bycatch and discards. Casablanca, Morocco, 26–28 May 2015. Food and Agriculture Organization of the United Nations, Rome

Faunce CH (2011) A comparison between industry and observer catch compositions within the Gulf of Alaska rockfish fishery. ICES J Mar Sci 68:1769–1777. https://doi.org/10.1093/icesjms/fsr093

Feekings J, Bartolino V, Madsen N, Catchpole T (2012) Fishery discards: factors affecting their variability within a demersal trawl fishery. PLoS ONE 7:e36409. https://doi.org/10.1371/journal.pone.0036409

Fox DS, Starr RM (1996) Comparison of commercial fishery and research catch data. Can J Fish Aquat Sci 53:2681–2694. https://doi.org/10.1139/f96-230

Gezelius SS (2006) Monitoring fishing mortality: compliance in Norwegian offshore fisheries. Mar Policy 30:462–469. https://doi.org/10.1016/j.marpol.2005.06.004

Gianuca D, Phillips RA, Townley S, Votier SC (2017) Global patterns of sex- and age-specific variation in seabird bycatch. Biol Conserv 205:60–76. https://doi.org/10.1016/j.biocon.2016.11.028

Gilman E (2015) Status of international monitroing and management of abandoned, lost and discarded fishing gear and ghost fishing. Mar Policy 60:225–239. https://doi.org/10.1016/j.marpol.2015.06.016

Gilman E, Suuronen P, Hall M, Kennelly S (2013) Causes and methods to estimate cryptic sources of fishing mortality. J Fish Biol 83:766–803. https://doi.org/10.1111/jfb.12148

Gilman E, Chaloupka M, Dagorn L et al (2019) Robbing Peter to pay Paul: replacing unintended cross-taxa conflicts with intentional tradeoffs by moving from piecemeal to integrated fisheries bycatch management. Rev Fish Biol Fish 29:93–123. https://doi.org/10.1007/s11160-019-09547-1

Gilman E, Perez Roda A, Huntington T et al (2020) Benchmarking global fisheries discards. Sci Rep 10:14017. https://doi.org/10.1038/s41598-020-71021-x

Gray CA, Kennelly SJ (2018) Bycatches of endangered, threatened and protected species in marine fisheries. Rev Fish Biol Fish 28:521–541. https://doi.org/10.1007/s11160-018-9520-7

Gullestad P, Blom G, Bakke G, Bogstad B (2015) The "Discard Ban Package": experiences in efforts to improve the exploitation patterns in Norwegian fisheries. Mar Policy 54:1–9. https://doi.org/10.1016/j.marpol.2014.09.025

Gullestad P, Abotnes AM, Bakke G et al (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. https://doi.org/10.1016/j.marpol.2016.11.032

Hall MA (1996) On bycatches. Rev Fish Biol Fish 6:319–352. https://doi.org/10.1007/BF00122585

Hammond TR, Trenkel VM (2005) Censored catch data in fisheries stock assessment. ICES J Mar Sci 62:1118–1130. https://doi.org/10.1016/j.icesjms.2005.04.015

Harrison XA, Donaldson L, Correa-Cano ME et al (2018) A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ 6:e4794. https://doi.org/10.7717/peerj.4794

Heath MR, Cook RM (2015) Hind-casting the quantity and composition of discards by mixed demersal fisheries in the North Sea. PLoS ONE 10:e0117078. https://doi.org/10.1371/journal.pone.0117078

Hoare D, Graham N, Schon PJ (2011) The Irish Sea data-enhancement project: comparison of self-sampling and national data-collection programmes-results and experiences. ICES J Mar Sci 68:1778–1784. https://doi.org/10.1093/icesjms/fsr100

Hylen A, Jacobsen JA (1987) Estimation of cod taken as by-catch in the Norwegian fishery for shrimp north of 69 °N. ICES CM 1987/G:34, Copenhagen

ICES (2000) Report of the study group on discard and by-catch information. ICES, ICES CM 2000/ACFM:11, Copenhagen

ICES (2007a) Report on the workshop on discard raising procedures. ICES WKDRP CM 2007/ACFM:06 Ref RMC PGCCDBS, San Sebastian

ICES (2007b) Report of the workshop on using fishers to sample catches (WKUFS). CM 2007/ACFM:24 Ref RMC PGCCDBS, Bergen

ICES (2020) Working group on methods for estimating discard survival. ICES Sci Rep 2:8. https://doi.org/10.17895/ices.pub.6003

James KM, Campell N, Viðarsson JR et al (2019) Tools and technologies for the monitoring, control and surveillance of unwanted catches. In: Uhlmann SS, Ulrich C, Kennelly SJ (eds) The European landing obligation: reducing discards in complex, multi-species and multi-jurisdictional fisheries. Springer, Cham, pp 363–382. https://doi.org/10.1007/978-3-030-03308-8

Johnson TR, van Densen WLT (2007) Benefits and organization of cooperative research for fisheries management. ICES J Mar Sci 64:834–840. https://doi.org/10.1093/icesjms/fsm014

Kaiser MJ, Clarke KR, Hinz H et al (2006) Global analysis of repsonse and recovery of benthic biota to fishing. Mar Ecol Prog Ser 311:1–14. https://doi.org/10.3354/meps311001

Karp WA, Breen M, Borges L et al (2019) Strategies used throughout the world to manage fisheries discards—lessons for implementation of the EU Landing Obligation. In: Uhlmann SS, Ulrich C, Kennelly SJ (eds) The European Landing Obligation: reducing discards in complex, multi-species and multi-jurisdictional fisheries. Springer, Cham, pp 3–26. https://doi.org/10.1007/978-3-030-03308-8

Kelleher K (2005) Discards in the world's fisheries. An update. FAO Fisheries Technical Paper. No. 470. FAO, Rome

Kennelly SJ (2020) Bycatch Beknown: methodology for jurisdictional reporting of fisheries discards—using Australia as a case study. Fish Fish 21:1046–1066. https://doi.org/10.1111/faf.12494

Kraan M, Uhlmann S, Steenbergen J et al (2013) The optimal process of self-sampling in fisheries: lessons learned in the Netherlands. J Fish Biol 83:963–973. https://doi.org/10.1111/jfb.12192

Liggins GW, Bradley MJ, Kennelly SJ (1997) Detection of bias in observer-based estimates of retained and discarded catches from a multi species trawl fishery. Fish Res 32:133–147. https://doi.org/10.1016/s0165-7836(97)00053-2

Lohr SL (2010) Sampling: design and analysis, 2nd edn. Chapman and Hall/CRC, Boca Raton

Lordan C, Ó Cuaig M, Graham N, Rihan D (2011) The ups and downs of working with industry to collect fishery-dependent data: the Irish experience. ICES J Mar Sci 68:1670–1678. https://doi.org/10.1093/icesjms/fsr115

Mangi SC, Dolder PJ, Catchpole TL et al (2013) Approaches to fully documented fisheries: practical issues and stakeholder perceptions. Fish Fish 16:426–452. https://doi.org/10.1111/faf.12065

Mangi SC, Kupschus S, Mackinson S et al (2018) Progress in designing and delivering effective fishing industry-science data collection in the UK. Fish Fish. https://doi.org/10.1111/faf.12279

Martin TG, Wintle BA, Rhodes JR et al (2005) Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecol Lett 8:1235–1246. https://doi.org/10.1111/j.1461-0248.2005.00826.x

Mayo RK, Lange AM, Murawski SA, et al (1981) A procedure for estimating rates of escapement and discard, based on research vessel bottom trawl survey catches. ICES CM 1981/G:62 Demersal Fish Committee

McBride MM, Fotland Å (1996) Estimation of unreported catch in a commercial trawl fishery. J Northwest Atl Fish Soc 18:31–41. https://doi.org/10.2960/J.v18.a2

National Research Council (2006) Review of recreational fisheries survey methods. The National Academies Press, Washington, DC. https://doi.org/10.17226/11616

Nedreaas K, Iversen S, Grethe HK (2015) Preliminary estimates of total removals by the Norwegian marine fisheries, 1950–2010. Working Paper #2015-94, Fisheries Centre, University of British Columbia, Vancouver

Needle CL, Dinsdale R, Buch TB et al (2014) Scottish science applications of remote electronic monitoring. ICES J Mar Sci 72:1214–1229. https://doi.org/10.1093/icesjms/fsu225

NMFS (2011) U.S. National Bycatch Report. U.S. Department of Commerce, NOAA Technical Memo NMFS-F/SPO-117E

NOU (2019) Framtidens fiskerikontroll. Norges ffentlige utredninger 2019, vol 21. Departementenes sikkerhets- og serviceorganisasjon, Oslo, p 254

Oliver S, Braccini M, Newman SJ, Harvey ES (2015) Global patterns in the bycatch of sharks and rays. Mar Policy 54:86–97. https://doi.org/10.1016/j.marpol.2014.12.017

Pálsson ÓK (2003) A length-based analysis of haddock discards in Icelandic fisheries. Fish Res 59:437–446. https://doi.org/10.1016/S0165-7836(02)00017-6

Pennington M (1983) Efficient estimators of abundance, for fish and plankton surveys. Biometrics 39:281–286. https://doi.org/10.2307/2530830

Pennino MG, Vilela R, Valeiras J, Bellido JM (2017) Discard management: a spatial multi-criteria approach. Mar Policy 77:144–151. https://doi.org/10.1016/j.marpol.2016.12.022

Pérez Roda MA, Gilman E, Huntington T et al (2019) A third assessment of global marine fisheries discards. FAO Fisheries and Aquaculture Technical Paper No. 633. FAO, Rome

Perretti CT, Deroba JJ, Legault CM (2020) Simulation testing methods for estimating unreported catch in a state-space stock assessment model. ICES J Mar Sci 77:911–920. https://doi.org/10.1093/icesjms/fsaa034

Pikitch EK, Santora C, Babcock EA et al (2004) Ecosystem-based fishery management. Science (80-) 305:346–347. https://doi.org/10.1126/science.1098222

Pitcher TJ, Watson R, Forrest R et al (2002) Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish Fish 3:317–339. https://doi.org/10.1046/j.1467-2979.2002.00093.x

Pitcher TJ, Kalikoski D, Pramod G, Short K (2009) Not honouring the code. Nature 457:658–659. https://doi.org/10.1038/457658a

Plet-hansen KS, Bastardie F, Ulrich C (2020) The value of commercial fish size distribution recorded at haul by haul compared to trip by trip. ICES J Mar Sci. https://doi.org/10.1093/icesjms/fsaa141

Punt AE, Smith DC, Tuck GN, Methot RD (2006) Including discard data in fisheries stock assessments: two case studies from south-eastern Australia. Fish Res 79:239–250. https://doi.org/10.1016/j.fishres.2006.04.007

Radford Z, Hyder K, Zarauz L et al (2018) The impact of marine recreational fishing on key fish stocks in European waters. PLoS ONE 13:e0201666. https://doi.org/10.1371/journal.pone.0201666

Rochet M-J, Trenkel VM (2005) Factors for the variability of discards: assumptions and field evidence. Can J Fish Aquat Sci 62:224–235. https://doi.org/10.1139/f04-185

Rudd MB, Branch TA (2017) Does unreported catch lead to overfishing? Fish Fish 18:313–323. https://doi.org/10.1111/faf.12181

Rue H, Martino S (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc 71:319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

Starr P (2010) Fisher-collected sampling data: lessons from the New Zealand experience. Mar Coast Fish 2:47–59. https://doi.org/10.1577/c08-030.1

Stock BC, Ward EJ, Thorson JT et al (2018) The utility of spatial model-based estimators of unobserved bycatch. ICES J Mar Sci 76:255–267. https://doi.org/10.1093/icesjms/fsy153

Stratoudakis Y, Fryer RJ, Cook RM (1998) Discarding practices for commercial gadoids in the North Sea. Can J Fish Aquat Sci 55:1632–1644. https://doi.org/10.1139/f98-029

Stratoudakis Y, Fryer RJ, Cook RM, Pierce GJ (1999) Fish discarded from Scottish demersal vessels: estimators of total discards and annual estimates for targeted gadoids. ICES J Mar Sci 56:592–605. https://doi.org/10.1006/jmsc.1999.0500

Suuronen P, Gilman E (2020) Monitoring and managing fisheries discards: new technologies and approaches. Mar Policy 115:103554. https://doi.org/10.1016/j.marpol.2019.103554

Tenningen M, Vold A, Olsen RE (2012) The response of herring to high crowding densities in purse-seines: survival and stress reaction. ICES J Mar Sci 69:1523–1531

Tenningen M, Pobitzer A, Handegard NO, de Jong K (2019) Estimating purse seine volume during capture: implications for fish densities and survival of released unwanted catches. ICES J Mar Sci 76:2481–2488. https://doi.org/10.1093/icesjms/fsz119

Thorson JT (2019) Guidance for decisions using the vector autoregressive spatio-temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fish Res 210:143–161. https://doi.org/10.1016/j.fishres.2018.10.013

Thorson JT, Scheuerell MD, Shelton AO et al (2015) Spatial factor analysis: a new tool for estimating joint species distributions and correlations in species range. Methods Ecol Evol 6:627–637. https://doi.org/10.1111/2041-210x.12359

Thorson JT, Ianelli JN, Larsen EA et al (2016) Joint dynamic species distribution models: a tool for community

ordination and spatio-temporal monitoring. Glob Ecol Biogeogr 25:1144–1158. https://doi.org/10.1111/geb.12464
Thorson JT, Barnett LAK, Anderson E (2017) Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. ICES J Mar Sci 74:1311–1321. https://doi.org/10.1093/icesjms/fsw193
Ulrich C, Olesen HJ, Bergsson H et al (2015) Discarding of cod in the Danish Fully Documented Fisheries trials. ICES J Mar Sci 72:1848–1860. https://doi.org/10.1093/icesjms/fsv028
Valdemarsen JW, Nakken O (2002) Utkast i norske fiskerier. Workshop om utkast i nordiske fiskerier, Sophienberg Slot 18–20 Nov 2002
van Helmond ATM, Mortensen LO, Plet-Hansen KS et al (2020) Electronic monitoring in fisheries: lessons from global experiences and future opportunities. Fish Fish 21:162–189. https://doi.org/10.1111/faf.12425
Veldhuizen LJL, Berentsen PBM, de Boer IJM et al (2018) Fish welfare in capture fisheries: a review of injuries and mortality. Fish Res 204:41–48. https://doi.org/10.1016/j.fishres.2018.02.001
Vigneau J (2006) Raising procedures for discards: sampling theory. ICES CM 2006/K:16, ICES, Copenhagen
Vlachavas A, Karaiskou N, Kokokiris L et al (2019) Using genetic methods for analysis of fish meals and feeds employed in Greek mariculture. Aquac Res 50:312–322. https://doi.org/10.1111/are.13900

PAPER II

ICES Journal of
Marine Science

ICES Journal of Marine Science (2022), 79(1), 76–87. https://doi.org/10.1093/icesjms/fsab242

Original Article

A simulation approach to assessing bias in a fisheries self-sampling programme

Thomas L Clegg, Edvin Fuglebakk, Kotaro Ono, Jon Helge Vølstad, and Kjell Nedreaas

Institute of Marine Research/Havforskningsinstituttet , Postboks 1870 Nordnes, NO-5817 Bergen, Norway

Corresponding author: tel: 55 23 85 00; e-mail: tom.clegg@hi.no

Clegg, T. L., Fuglebakk, E., Ono, K., Vølstad, J. H., and Nedreaas, K. A simulation approach to assessing bias in a fisheries self-sampling programme. – ICES Journal of Marine Science, 79: 76–87.

Received 5 October 2021; revised 15 November 2021; accepted 16 November 2021; advance access publication 5 December 2021.

The hierarchical structure and non-probabilistic sampling in fisher self-sampling programmes makes it difficult to evaluate biases in total catch estimates. While so, it is possible to evaluate bias in the reported component of catches, which can then be used to infer likely bias in total catches. We assessed bias in the reported component of catches for 18 species in the Barents Sea trawl and longline fisheries by simulating 2000 realizations of the Norwegian Reference Fleet sampling programme using the mandatory catch reporting system, then for each realization we estimated fleet-wide catches using simple design-based estimators and quantified bias. We then inserted variations (e.g. simple random and systematic sampling) at different levels of the sampling design (sampling frame, vessel, and operation) to identify important factors and trends affecting bias in reported catches. We found that whilst current sampling procedures for fishing operations were not biased, non-probabilistic vessel sampling resulted in bias for some species. However, we concluded this was typically within the bounds of expected variation from probabilistic sampling. Our results highlight the risk of applying these simple estimators to all species. We recommend that future estimates of total catches consider alternative estimators and more conservative estimates of uncertainty where necessary.

Keywords: design-based, hierarchical sampling, random forest, reference fleet, self-sampling.

Introduction

Self-sampling by fishers is emerging as an effective method of collecting data at sea (Starr, 2010; Lordan *et al.*, 2011; Roman *et al.*, 2011; Kraan *et al.*, 2013; Bell *et al.*, 2017). An extension of this approach is a reference fleet, defined as a group of active fishing vessels with an enhanced data collection role requiring training and support (Mangi *et al.*, 2013). Reference fleets can reduce the logistics and costs of data collection compared with observer programmes (Mangi *et al.*, 2013; Suuronen and Gilman, 2020), and improve relationships with the fishing industry through participatory research and a two-way communication channel (Kraan *et al.*, 2013).

As a relatively new approach, data collection through self-sampling has received more scrutiny than other more established methods such as scientific observers (Kraan *et al.*, 2013; Mangi *et al.*, 2013) towards the representativeness of samples for making inferences about the wider fleets it covers. Multi-stage sampling of catches in fisheries results in complex, hierarchical data. At each stage, there are chances that bias is introduced in either a self-sampling or independent observer programme (ICES, 2008; Table 1), reducing the representativeness of samples.

The bias of an estimate is the degree to which the expected value, obtained through repeated sampling and estimation, differs from the true value (Jessen, 1978). The reality of fisheries sampling makes it difficult to assess the bias of total catch estimates as the true value is not known for comparison. Under a landing obligation, such as those implemented in Norway and the EU, reported catches do not always accurately reflect total extractions from the fishery. Catches may go unreported due to illegal discarding, intentional misreporting, or because of low resolution of reporting for some species or species identification errors (Pitcher *et al.*, 2002). However, no single study can address all aspects of bias. For example, comparing

Table 1. Potential sources of bias in a fishery sampling programme.

Aspect of sampling design	Potential sources of bias
Sampling frame	• A poorly defined sampling frame will affect representativeness of samples. • Insufficient data to define the sampling frame risks out-of-frame samples being included.
Sampling units	
1) Vessel	• Opt-in participation may bias towards more compliant vessels (rejection effect). • Type and size of compensation may influence motivation to participate. • Mandatory participation may influence fishing strategy if 100% of fishing activity is not observed (Liggins *et al.*, 1997; Benoît and Allard, 2009; Snyder and Erbaugh, 2020).
2) Fishing operation	• Human selection of sampled fishing operations may favour convenience (e.g. sampling smaller hauls to reduce time spent sampling) and introduce bias.
3) Catches	• Gear characteristics, variable habitat, and fishing strategy will likely result in a non-random distribution of fish throughout the catch operation, which if sampled opportunistically would produce biased observations (e.g. sampling from the first available portion of the catch). • Intentional manipulation of catch data, possibly where results could have a large impact on management or policy decisions (e.g. misreporting catches of protected species or species with limited quota). • Poor sampling techniques, either through inadequate training (e.g. species identification, misuse of equipment) or lack of time during catch processing.
Estimator	• Choice of estimator must be upheld by the relevant assumptions

self-sampling data with independent observations of known reliability (Faunce, 2011; Roman *et al.*, 2011) will only address measurement error such as under-reporting of sensitive species. Similarly, comparing multiple estimators or sampling designs may address biases in those specific aspects (Diamond, 2003; Cahalan *et al.*, 2015; Cahalan and Faunce, 2020), but prior knowledge of potential biases will help when defining the candidate estimators.

Total catches can be broken down into the portion reported to the authorities (in daily logbooks or landing reports), and the portion that is not reported but does occur. The unreported component is especially problematic because it is difficult, if not impossible to quantify biases related to them (Ainsworth and Pitcher, 2005). However, we can reach a better understanding of biases affecting estimates of total catches by focusing on just the reported component. Mandatory catch reporting acts as a census of fishing effort and the reported component of total catches, so it can be utilized to explore such biases, assuming that biases affecting reported catches are also likely to affect total catches (Liggins *et al.*, 1997).

The Institute of Marine Research recruits vessels and maintains the Norwegian Reference Fleet, a fisheries self-sampling project, which tasks participating vessels with regularly gathering of data on fishing operations during normal fishing activity (Clegg and Williams, 2020). An independent evaluation of the Norwegian Reference Fleet by Bowering *et al.* (2011) concluded that based on expert judgement and limited analyses, the sampling programme is representative of the wider fleets it covers. However, Bowering *et al.* (2011) concluded that focused analyses are needed to evaluate representativeness of individual segments of the Norwegian Reference Fleet.

This study aims to understand the representativeness of the Norwegian Reference Fleet sampling design by identifying biases that are likely to affect estimates of total catches. We addressed this by focusing on the reported component of total catches, for which we have census data in the form of mandatory daily catch logbooks. We simulated Norwegian Reference Fleet sampling design in the Barents Sea trawl and longline fisheries between 2012 and 2018. We then estimated fleet-wide catches based on the reports in the sample and quantified bias through comparison with observed reported catches and used random forest models to understand which variables are important for explaining variations in bias. Assuming biases in our estimates of reported catches will likely affect estimates of total catches, we discuss the results in the context of total catches to suggest ways in which the biases can be reduced or mitigated.

Data

Case study fisheries

Our case study is focused on a portion of the Barents Sea bottom trawl and longline fisheries, defined as vessels with overall length (LOA) greater than 28 m using bottom trawl or longline fishing gears to target demersal fish species in the statistical areas highlighted in Figure 1. Area 24 only includes the trawl fishery, as the longline fishery does not extend into this statistical area. In the trawl fishery, a fishing operation is defined as a haul, whilst in the longline fishery it is defined as all hooks hauled from all longlines in a calendar day.

Both fisheries occur all year round, peaking between November and January. However, statistical areas 23 and 24 are typically inaccessible between January and April due to sea ice. Vessels predominantly target cod (*Gadus morhua*) and to a lesser extent haddock (*Melanogrammus aeglefinus*), but also infrequently target saithe (*Pollachius virens*), tusk (*Brosme brosme*), Greenland halibut (*Reinhardtius hippoglossoides*), and beaked redfish (*Sebastes mentella*).

The northern prawn (*Pandalus borealis*) trawl fishery and pelagic trawl fisheries (mainly capelin, *Mallotus villosus*) carry a small risk of being included in this study because of overlap in space and time and possible erroneous reporting of trawl gear codes. However, these fisheries are easily identifiable due to very high selectivity, allowing fishing operations to be removed if the dominant species was not a demersal fish.

For each year, we post-stratified samples based on a combination of statistical area (Figure 1) and season (winter: January–April; summer: May–August; and autumn: September–December). Three seasons reflect the seasonality of the Barents Sea, such as ice cover restricting access to areas 23 and 24 (Figure 1) in winter. Furthermore, there were insufficient data to estimate fleet-wide reported catches on a monthly timescale. We only estimated reported catches in strata with three or more fishing operations sampled, as unsam-

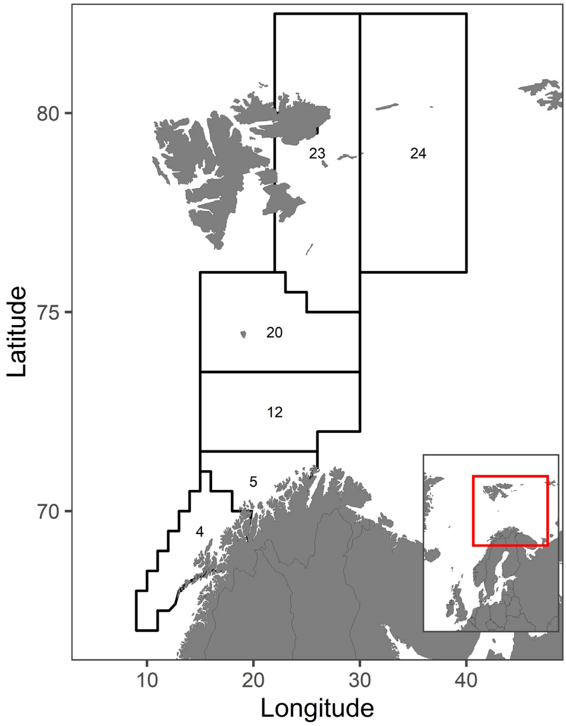


Figure 1. Statistical areas in the Barents Sea trawl and longline fisheries as defined by the Norwegian Directorate of Fisheries. Area 24 is excluded from the longline fishery due to negligible fishing activity.

pled strata require imputation methods that are often subjective and so introduce new biases that are difficult to quantify (Stratoudakis *et al.*, 1999; Lohr, 2010).

Table 2 provides a full list of species included in this study. Broad species groups were removed from the study (e.g. unidentified flatfish) to avoid double counting. An exception to this rule was skates and rays for which species identification is notoriously difficult and are typically grouped in the reported catches. Species observed in less than 1% of fishing operations in any given year were removed from the study. Extremely rare species typically require more complex modelling approaches for estimating total catches, so the relevance of bias using design-based estimators is not of relevance to this study.

Norwegian reference fleet

The Norwegian Reference Fleet project is a trust-based collaboration between fishers and scientists to improve data for input into stock assessments and provide data on bycatches and discards. Our selected fisheries in this study are prioritized for the Norwegian Reference Fleet, meaning that active participation is required in the contract to ensure representativeness and sufficient data for stock assessments.

Table 2. Species included in this study. Asterisks mark species included only in the longline fishery.

Common name	Scientific name	FAO code
Atlantic wolffish	*Anarhichas lupus*	CAA
Northern wolffish*	*Anarhichas denticulatus*	CAB
Spotted wolffish	*Anarhichas minor*	CAS
Ratfish*	*Chimaera monstrosa*	CMO
Atlantic cod	*Gadus morhua*	COD
Greater forkbeard*	*Phycis blennoides*	GFB
Greenland halibut	*Reinhardtius hippoglossoides*	GHL
Haddock	*Melanogrammus aeglefinus*	HAD
Atlantic halibut	*Hippoglossus hippoglossus*	HAL
Common ling	*Molva molva*	LIN
Monkfish*	*Lophius piscatorius*	MON
Saithe	*Pollachius virens*	POK
Skates and rays*	*Rajidae*	RAJ
Beaked redfish	*Sebastes mentella*	REB
Golden redfish	*Sebastes norvegicus*	REG
Roughhead grenadier*	*Macrourus berglax*	RHG
Lesser redfish*	*Sebastes viviparus*	SFV
Tusk	*Brosme brosme*	USK

Vessel owners can apply to participate in the Norwegian Reference Fleet through a publicly open tender process, as mandated by law. This process involves a public announcement of a paid contract for a vessel to sample catches from their normal fishing activity over a 4-year period. Each tender specifies a list of mandatory and desired requirements, which aim to recruit a typical vessel in the defined category. Applicants must provide evidence that they meet these requirements to be included in the selection process. Eligible applications are then assessed by a panel to evaluate the desired requirements. If after this evaluation there are multiple eligible vessels, then the contract is awarded randomly.

The vessel categories relevant to this study are the trawl and longline categories where the tender requires that vessels must be greater than 39 m and 35 m LOA, respectively, and hold permit and quota to fish various demersal species in the Barents Sea fisheries. Over the study period, around 14% of all vessels in the study fisheries have participated in the Norwegian Reference Fleet (longline: 8/55 vessels; trawl: 6/42 vessels).

Each vessel has designated crew who are given training on the sampling protocol and species identification. These crew sample total catches systematically, such that one fishing operation is sampled every 2 days (1-in-2 systematic sample; Lohr, 2010). The starting day for systematic sampling is selected randomly for each trip by the crew or skipper. For trawl vessels, the fishing operation is defined as a single haul. A haul is randomly selected from all those planned on the sampling day. On longline vessels, for which fishing operation is defined as a calendar day, it is relatively easier to subsample the catch in a fishing operation. All specimens are recorded from a subsample of consecutive hooks, spanning the start, middle, and end of longlines that the crew or skipper deems representative of the catch composition for that day.

Daily logbooks

Norwegian law requires all vessels in the Barents Sea fisheries at or above 15 m LOA to report the weight of each species caught in every fishing operation through an electronic reporting system (ERS). Each entry contains an estimated total live weight for each species, alongside the time and location of the fishing operation. This at-sea reporting of total catches may be biased downwards if fish are discarded during processing or processed catches are misreported (i.e. illegally landed). Official catch statistics are reported as round weight (live weight when removed from the sea), but on factory vessels, all catch reporting is done post-production. Therefore, product weights are converted back to round weight using official conversion factors for each product (Norwegian Directorate of Fisheries, 2021). This weight conversion has negligible impact on our study as factors are applied consistently across all vessels. Although reported weights are estimated at sea, various regulations ensure that weights match those officially declared in sales notes when weighed and sold on land (10% tolerance) to verify the accuracy of reported catches (Gezelius, 2006; Gullestad *et al.*, 2015). Whilst catches are weighed more accurately in sales notes, they offer insufficient data resolution for our analysis because they are a summary of all fishing activity in each trip, which spans multiple statistical areas over a period of weeks. We, therefore, can view the ERS data as a census of true reported catches in the fisheries for the purposes of this study.

We extracted ERS logbook data, provided by the Norwegian Directorate of Fisheries, between 2012 and 2018 from all commercial fishing operations (longline: 21807; trawl: 109801) in the two case study fisheries. We removed 17 entries from trawl vessels for which trawl duration could not be calculated because erroneous start and stop times were reported. We also removed one anomalous entry where a trawl vessel reported 40 tonnes of fish for 1 min of fishing time.

Methods

In order to identify biases that are likely to affect estimates of unreported catches in the Barents Sea trawl and longline fisheries between 2012 and 2018, we performed a simulation study using the ERS logbook data on reported catches. By simulating the Norwegian Reference Fleet sampling design using data on reported catches, we can generate estimates of the reported component of catches that can be directly compared to the true values, such that biases can be quantified. Results from these simulations can then be applied to improve estimates of total catches (i.e. reported + unreported) using the real data generated by the Norwegian Reference Fleet sampling programme mimicking the Norwegian Reference Fleet sampling design and various other designs (e.g. simple random), and quantified bias through comparison of the estimated fleet-wide catches from the simulated sample with observed reported catches.

Simulating sampling designs

The simulation framework consists of three components which we manipulated: the sampling frame, vessel sampling, and fishing operation sampling. Simulating these in a fully-crossed design (Figure 2) ensured balanced groups for the statistical analysis (Boulesteix *et al.*, 2012).

We assessed if the Norwegian Reference Fleet is representative of the entire fishery (> 28 m LOA) by first estimating total reported catches of vessels within the length ranges defined in Norwegian Reference Fleet tender specifications (sampling frame: *LIM*; Figure 2) and compared it to an estimation of reported catches from all vessels in the study fishery (sampling frame: *ALL*).

To evaluate the bias incurred from the non-probabilistic selection of vessels in the current Norwegian Reference Fleet (vessel: *RF*; Figure 2), we also performed two probabilistic selections of vessels for comparison. First, a simple random sample (vessel: SRS_V) of vessels within the vessel length class requirements of the Norwegian Reference Fleet tender specifications, and second, a weighted random sample (vessel: *WRS*) where vessels were selected with a probability proportional to the fishing effort (fishing days) in the previous year.

Finally, we simulated the 1-in-2 systematic sampling protocols by the Norwegian Reference Fleet (fishing operation: *SYS*; Figure 2), then compared it to a simple random sample (fishing operation: SRS_{FO}) from all fishing operations by sampled vessels in each stratum, with a sample size equal to that of the post-stratified systematic sample for each simulation. Systematic sampling is expected to be the equivalent of simple random sampling but was nevertheless included for confirmation.

Estimation procedure

We chose two conventional design-based estimators [Equations (1) and (2)] that are currently used for estimating discards and bycatches in Norwegian fisheries (Bærum *et al.*, 2019; Berg and Nedreaas, 2020; Moan *et al.*, 2020). These simple estimators assume that samples are randomly selected from all fishing operations in

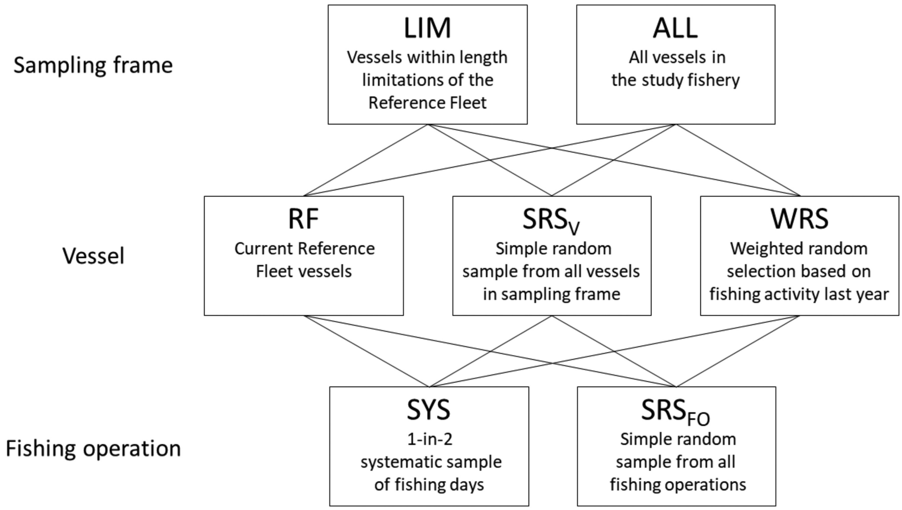


Figure 2. Schematic diagram of the simulations showing the three stages and fully-crossed design.

each stratum, effectively ignoring variations in catches between vessels. Despite the hierarchical sampling design of the Norwegian Reference Fleet, it is not yet clear how a multi-stage estimator should be defined due to the non-probabilistic selection of vessels and a lack of understanding of which levels of the sampling design contribute the most to variations in total catches. Therefore, for the purposes of simplicity, we ignored the hierarchical sampling design and applied commonly used simple estimators.

For species i in year j, total reported catches ($\hat{Y}_{i,\ j}$) in sampled strata were estimated using the stratified unit estimator [Equation (1)] for the longline fishery, and the stratified ratio estimator [Equation (2)] for the trawl fishery:

$$\hat{Y}_{i,j} = \sum_{k=1}^{K} \frac{N_{j,k}}{n_{j,k}} y_{i,j,k}, \tag{1}$$

$$\hat{Y}_{i,j} = \sum_{k=1}^{K} \frac{y_{i,j,k}}{x_{j,k}} X_{j,k}, \tag{2}$$

where $k =$ stratum, $y =$ weight of reported catches sampled, $n =$ number of fishing operations sampled, $N =$ total number of fishing operations in population, $x =$ sampled trawl duration, and $X =$ total trawl duration in population. Note that when simulating the Norwegian Reference Fleet sampling design for longline vessels, we do not need to account the subsampling of hooks (which is done for Norwegian Reference Fleet sampling), as catches for the entire fishing operations are reported in the ERS logbooks.

The accuracy of the ratio estimator [Equation (2)] is partially dependent on the correlation between catch weight (y) and trawl duration (x). We, therefore, calculated the Spearman's rank correlation coefficient between catch weight and trawl duration for each species (Rochet and Trenkel, 2005; Lohr, 2010).

The re-sampling and estimation processes were repeated 2000 times (S = 2000), and the relative error (RE) of $\hat{Y}_{i,\ j}$ calculated as a measure of bias using:

$$RE_{i,j} = \frac{1}{S} \sum_{s=1}^{S} \frac{\hat{Y}_{i,j,s} - Y_{i,j,s}}{Y_{i,j,s}}, \tag{3}$$

where $Y_{i,\ j,\ s}$ is the observed annual catches of species i in sampled strata in year j and in simulation s.

Systematic sampling of fishing operations for each vessel will result in a different sample size for every realization of the sampling protocol. As the starting day for systematic sampling is selected for each vessel per simulation, the number of fishing operations in each stratum will vary slightly depending on where the sampling days fall, which is in turn proportional to the number of vessels in that year. For trawl vessels, this is further affected by which haul is selected for sampling on each day. These variations were not controlled as they reflect the true variations that arise from sampling protocols.

Modelling the sources of bias

We used random forests (Breiman, 2001) as a regression method to explore which aspects of the sampling process are causing bias in estimations of reported catches. For a detailed explanation and guidance for random forest models in various biological and ecological contexts, we recommend Cutler *et al.* (2007), Boulesteix *et al.* (2012), Fox *et al.* (2017), and Siders *et al.* (2020).

First and foremost, we chose random forests for a unique and novel feature: variable importance (Genuer *et al.*, 2009). Fitting a random forest model involves repeatedly sampling observations and explanatory variables from the original dataset. Those observations not included in model fitting (known as "out-of-bag" samples) can be used to test the accuracy of model prediction (comparable to cross validation in Generalized linear models). However, random forests can also examine how the exclusion of each explanatory variable affected the accuracy of prediction. This overall decrease in accuracy can be used as a measure of importance for each

Table 3. Variables considered in the random forest models.

Variable	Type	Description
Vessel	Categorical	Vessel sampling design simulated†
Fishing operation	Categorical	Fishing operation sampling design simulated†
Sampling frame	Categorical	Population from which vessels are sampled†
Species	Categorical	Species or species group (Table 2)
Year	Categorical	Year (2012–2018)
Encounter rate	Continuous	Proportion of fishing operations where the species was observed (mean across all simulations).
Sample size*	Continuous	Number of fishing operations sampled annually (mean across all simulations)
Correlation coefficient*	Continuous	Spearman's rank correlation coefficient between catch weight and trawl duration for each species and year

†See Figure 2 for categories.
*Only used in trawl fishery model as it is only relevant to the bias of a ratio estimator.

explanatory variable. For example, if there was no change in model accuracy when a variable was not included, then that variable was not important in explaining any variations in bias.

Random forest methods also suit the exploratory nature of our study, mitigating against the bad practice of "data dredging" (Burnham and Anderson, 2002). There is no iterative variable selection and testing process, meaning all potential variables are included in a single model. Similarly, random forests automatically capture non-linear relationships and complex interactions between variables without the need for prior specification.

To account for possible biases arising from differences in scale and type of predictor variables, we used a class of decision trees called conditional inference trees to build the random forest (Hothorn *et al.*, 2006; Strobl *et al.*, 2007). All models were fitted using the R package *party* (version 1.3–7; Hothorn *et al.*, 2006; Strobl *et al.*, 2007, 2008) with "*cforest_unbiased*" convenience controls. Importance was estimated conditionally, which accounts for correlated variables and interactions (Strobl *et al.*, 2008, 2009). We also explored the relationship between bias and each explanatory variable using partial dependence plots (Friedman, 2001), calculated using the R package *edarf* (version 1.1.1; Jones and Linder, 2016).

In addition to the various elements of sampling design we simulated (Figure 2), we also included five other predictor variables in the random forest analysis: species, encounter rate, year, sample size, and correlation coefficient (Table 3). Sample size and correlation coefficient were only included in the trawl fishery model where we used a ratio estimator, for bias is a function of these two variables. As an alternative to encounter rate, we considered total annual reported catches of all species. However, encounter rate was preferred as a better indication of the rarity of species and was more relevant to the sampling data than the population data.

We fitted a random forest model to each fishery independently. Model tuning was steered by two hyperparameters: number of trees (*ntree*) and number of randomly sampled variables at each split (*mtry*), which were optimized using a simplified grid search to minimize the out-of-bag mean square error (MSE). We tested all possible *mtry* values (longline: 1–6; trawl: 1–8 variables), and five *ntree* values spaced evenly between 50 and 1000. Results may be sensitive to the random seed, so we compared variable importance outputs from five initial runs with new random seeds. Any substantial changes in results indicates instability, meaning more trees should be added. Tuning resulted in both models being fitted using $mtry = 4$ and $ntree = 500$ (Supplementary Table S1).

Additional analyses

The current selection of vessels in the Norwegian Reference Fleet is the only possible realization of the expert judgement selection process. Comparatively, we have simulated a large number of realizations using a simple random sample of vessels. We assume the non-probabilistic, expert judgement selection behaves like a simple random sample to be able to use design-based estimators. For this assumption to be upheld, we would expect that the estimate based on expert judgement selection of vessels would lie within the distribution of estimates based on a simple random selection. We evaluated this using z-scores for each estimate of relative error, taking the mean and standard deviation from the distribution of relative errors from simulations using a simple random sample of vessels. In addition to this, we made pairwise comparisons of accuracy using the different vessel sampling methodologies for individual simulations to determine how often a simple random sample out-performs the expert judgement selection.

Results

Amongst the three components of the catch sampling design (Figure 2), vessel sampling method was the most important contributor of bias when estimating the reported component of total catches in both the longline and trawl fisheries (Figure 3). With the estimators used here, the current realization of the Norwegian Reference Fleet vessel selection procedure tends to overestimate catches when averaged across all species (Figure 4a). However, this summary statistic does not account for the important variations in bias across years and species (Figure 4b and c).

Figure 5 illustrates to what extent the expert judgement selection in the current Norwegian Reference Fleet behaves like simple random sampling. More than half of the annual species catch estimates using a non-random vessel selection were within one standard deviation of the mean from a simple random sample (z-score < 1; longline: 56%; and trawl: 77%), with few estimates being outside the 95% confidence interval (z-score > 1.96; longline: 12%; and trawl: 4%). However, the distribution of z-scores was skewed to the right (Figure 5), confirming the tendency for this non-random sample of vessels to result in overestimation in reported catches of individual species.

As the sampling frame was not important in explaining the variations in bias when estimating the reported component of catches for individual species (Figure 3), it suggests that the vessel length

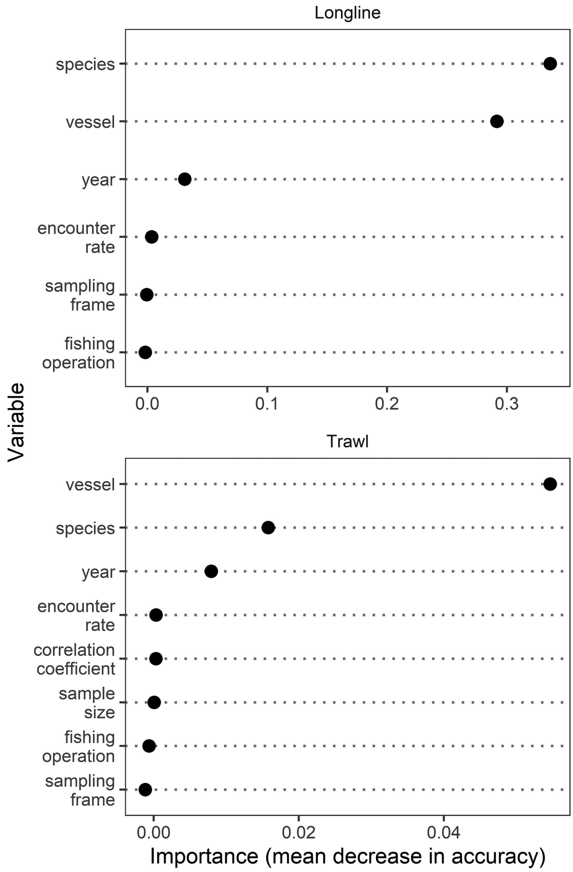


Figure 3. Importance of variables for explaining variations in relative error when predicting the reported component of total catches. Measures of conditional importance of variables estimated using random forest models. Note different scales on *x*-axes.

requirements in the Norwegian Reference Fleet do not affect the representativeness in relation to all vessels in the fishery (> 28 m LOA). Similarly, the simulated methodologies for sampling fishing operations were of negligible importance when estimating total reported catches (Figure 3), confirming that systematic sampling of fishing operations is the equivalent of simple random sampling.

Variations in relative error of estimated reported catches between species was of high importance in both fisheries and was the most important variable in the longline fishery (Figure 3). These variations across species were large in both fisheries but was most extreme in the longline fishery (Figure 4c). For many species, the probabilistic sample of vessels still resulted in estimation bias, which further highlights that the weak assumptions in our chosen estimators (i.e. not applying multi-stage estimators, and poor correlation between catches and trawl duration) resulted in bias due to species-specific factors. Total reported catches were underestimated only in a small number of cases, most notably beaked redfish (REB) and skates and rays (RAJ) in the longline fishery.

In the trawl fishery, biases were consistently largest across species when using the non-random selection of vessels (Figure 4c). The ratio estimator [Equation (2)] is known to be biased when the correlation between catches and trawl duration is low, and with a smaller sample size (Lohr, 2010). However, the random forest model indicated both these factors were of low importance when explaining variations in bias (Figure 3). Given the small variations in these explanatory variables (Supplementary Figure S1), this does not provide insights into how bias would improve with higher correlation coefficients or larger sample size.

Despite accounting for many other variables that could explain differences in relative error, there were still annual variations affecting bias of reported catch estimates across all species, albeit weak (Figure 3). These variations were more extreme for non-random Norwegian Reference Fleet vessel sampling (Figure 4b) than for probabilistic sampling of vessels.

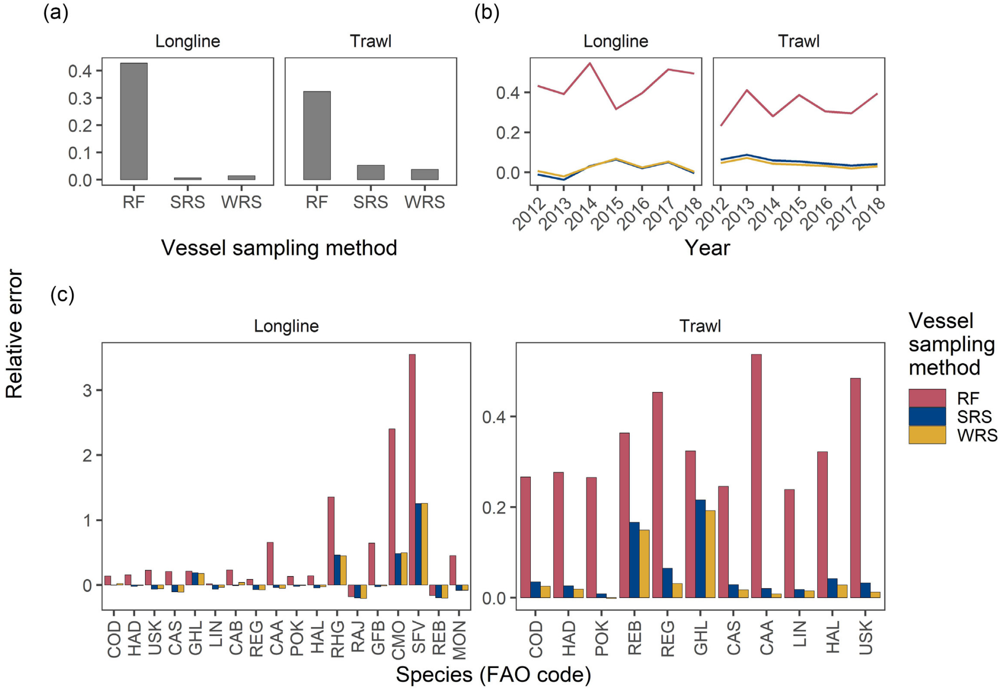


Figure 4. Partial dependence plots of the most important variables in the random forest models, showing the marginal effect of variables on relative error. (a) Vessel selection (RF = current Norwegian Reference Fleet vessels (fixed); SRS = simple random sample; and WRS: weighted random sample). (b) and (c) include interactions between the three vessel sampling methodologies. Note different scales on *y*-axes in panel c. Species ordered by descending total reported catches for each fishery. Species names are listed in Table 2.

Discussion

Through a simulation approach using the reported component of total catches, we have identified aspects of the Norwegian Reference Fleet vessel selection and catch sampling design that will likely affect biases in estimates of total catches for a range of species in the Barents Sea trawl and longline fisheries. We have identified that the vessel selection process for the Norwegian Reference Fleet resulted in an overestimation of the reported component of total catches for many species, indicating possible biases in similar estimations of unreported catches. However, we also found biases in reported catch estimates using probabilistic sampling, which indicates the hierarchical structure of Norwegian Reference Fleet programme's sampling design does not meet the assumptions of the simple estimators used here. Furthermore, important variations between species indicate that the suitability of our chosen estimators is highly dependent on species-specific factors, suggesting the consideration of alternative estimators.

Vessel selection process

There is potential for the vessel selection process to result in bias due to practical difficulties in maintaining a reliable at-sea sampling programme such as issues regarding participation (Vølstad and Fogarty, 2006; Benoît and Allard, 2009; Kraan *et al.*, 2013) and budget restrictions (Borges *et al.*, 2004). For a simple random selection of vessels, one realization could risk selecting vessels with very different fishing strategies relative to the wider fleet. Furthermore, randomly selected vessels may be rarely active in the fishery, yielding little data for a large investment in equipment and training. A weighted random sample could reduce those risks but could also introduce further complications. For example, a vessel that is consistently active in the fishery could spend long periods in the harbour for repairs or refurbishments, reducing its selection probability for the following year. Probabilistic sampling implies that all vessels are willing to participate if selected, which is unrealistic if not impossible, as current laws require that all vessels must have the opportunity to apply to an open tender, rather than being selected and requested to participate.. Mandatory participation can function with independent samplers (Ewell *et al.*, 2020), but self-sampling requires large amounts of time and effort alongside normal fishing activity. If self-sampling was mandatory and without compensation, then we would expect trust to deteriorate between scientists and fishers and a reduction in data quality (Jacobsen *et al.*, 2012; Mangi *et al.*, 2016). Our results found that for the majority of cases, an estimate using expert judgement selection of vessels for the Norwegian Reference Fleet was within the range of expected estimates using a simple random selection of vessels (Figure 5). However, there is a tendency when using Norwegian Reference Fleet to overestimate reported catches, which can be significantly large for certain species.

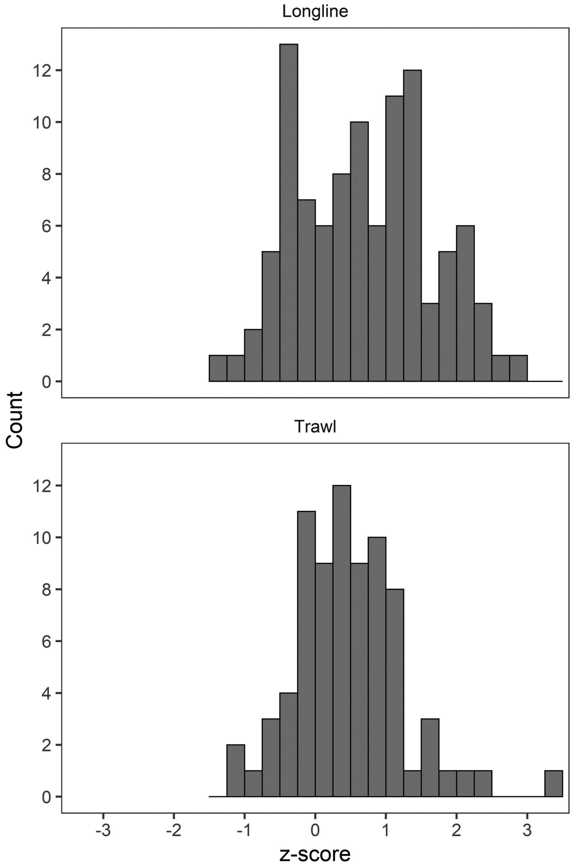


Figure 5. The range of z-scores for bias in annual estimates of reported catches for species using an expert judgement selection of vessels. A z-score is the number of standard deviations from the mean of simulated estimates using a simple random sample of vessels (i.e. where an expert judgement selection lies within a distribution of simple random samples). Each z-score represents an estimate of annual reported catches for one species.

Supplementary Figures S2 and S3 in the Supplementary materials provide comparisons of fishing activity between Norwegian Reference Fleet vessels and the wider fleet to suggest reasons for the tendency of Norwegian Reference Fleet sampling to overestimate catches. Norwegian Reference Fleet vessels have higher catch rates and different fishing strategies compared to the wider fleet. These differences could be caused by higher engine power and annual quotas, both of which will affect overall catch composition, and therefore, unreported catches. Furthermore, in selecting a small sample of vessels that are to be fixed for several years, the selection cannot be optimized to capture the tails of statistical distributions that may be derived from these vessels or their catches. It is reasonable to expect this to lead to some underestimation of variability of total catches, and some over- or under-estimation of total catches when the distribution of total catches is not symmetric.

Although a devoted study is required to provide evidence for the differences in catch rates, we can still consider ways to improve incentives for participating in the Norwegian Reference Fleet. A larger number of applications would improve the outcome of expert selection, and in the situation where similar vessels are short-listed, then random selection could be applied more effectively.

Variations in bias across species

We cannot discuss the biases in estimates for individual species in this exploratory study, but we can nevertheless discuss generalizations. Each species and stratum can be viewed as a domain (subpopulation), as these properties are unknown before the sample is taken (Lohr, 2010). Whilst there are methods to approximate the sampling probabilities of fishing operations (i.e. systematic sampling for the Norwegian Reference Fleet), a single sampling programme covering all species will not adequately address the sampling demands for all species (domains), such as spatial and temporal variations (Stock *et al.*, 2020), catchability (Rochet and Trenkel, 2005), and sorting or reporting behaviour (Pitcher *et al.*, 2002). However, our results suggest that excessively large overestimations are limited to a small number of rare or non-commercial bycatch species (Figure 4). The practical consequences of non-probabilistic sampling may vary for any specific parameters observed, and across species. For example, a small spatial bias can be very significant for very localized populations if not adequately accounted for in the chosen estimator (Cosandey-Godin *et al.*, 2014).

Choice of estimator

The interpretation of our results must be viewed through the lens of the estimators we applied. We applied simple estimators as they are currently the standard practice for estimating discards and bycatches using data from the Norwegian Reference Fleet. However, reported catches tended to be overestimated even when vessels were sampled probabilistically. The magnitude of this overestimation when using probabilistic sampling is small relative to estimates using observations from Norwegian Reference Fleet vessels. However, it is still important to consider why this overestimation is occurring, and why it is larger in the trawl fishery.

The complex hierarchical structure of fisheries sampling is often ignored to simplify estimations (Nelson, 2014), either based on assumption or from evidence that sampling levels do not contribute significantly to the total variance (Tamsett *et al.*, 1999). However, accounting for the variations between sampled vessels can improve both the accuracy and precision of estimates (Moan *et al.*, 2020; Fernandes *et al.*, 2021). Our results, therefore, indicate that ignoring the hierarchical structure of Norwegian Reference Fleet sampling could result in biased estimates in total catches. This bias may be mitigated by accounting for the hierarchical structure of sampling, namely the sampling of vessels. Adequately handling the two-stage sampling design will also allow for variance estimates that may both explain part of the perceived bias and put the bias in perspective. We found small biases in our estimations in the trawl fishery, even when using probabilistic sampling of both vessels and fishing operations. The accuracy of ratio estimators is influenced by both the sample size and the strength of correlation between catches and fishing effort (Lohr, 2010). We applied the ratio estimator as this is typically employed in studies which assume a relationship between catches and fishing effort, despite evidence of little to no relationship in many cases (Diamond, 2003; Rochet and Trenkel, 2005; Cahalan *et al.*, 2015), given that other influential variables have not been accounted for, such as vessel characteristics, or on a finer spatio-temporal scale, such as within-stratum variations or variations across trips. The random forest analysis indicated that both the correlation coefficient and sample size were not important for explaining variations in biases when estimating the reported component of catches (Figure 3). However, this may be because of the limited range in values for the correlation coefficient and sample size (Supplementary Figure S3). Our random forest model was built to explore variations in biases, so whilst some variables may not have been deemed important in explaining variations in bias, it suggests that the ratio estimator was consistently biased across all species and years.

Limitations

We must also acknowledge additional biases related to total catch sampling that could not be addressed in this study. For example, official reporting of catches is enforced through inspection and carries a risk of prosecution if inaccurate. Conversely, voluntary recording of total catches may risk under-reporting sensitive portions of the catch (measurement biases; Table 1). However, based upon the data sent by Norwegian Reference Fleet vessels, the constant dialogue between scientists and fishers, payment for sampling, and supervision, we are generally confident that data collectors in the Norwegian Reference Fleet report honestly. The willingness to report sensitive information is evident in studies, which have used Norwegian Reference Fleet data to estimate "unsustainable" and "concerning" levels of bycatches of sensitive species such as seabirds (Fangel *et al.*, 2015; Bærum *et al.*, 2019) and harbour porpoise (*Phocoena phocoena*; Moan *et al.*, 2020), which are entirely absent from the official reporting framework. Nevertheless, comparisons of Norwegian Reference Fleet data with a data source of known reliability (e.g. independent observer, remote electronic monitoring; Liggins *et al.*, 1997; Faunce, 2011) could test this assumption and improve reliability from a statistical perspective (Kraan *et al.*, 2013).

This simulation study focused on the selection of vessels and hauls which was limited by the available data from mandatory reporting of catches. However, it is also important to consider biases in the biological sampling of catches which would affect size-based estimates of unreported catches. For example, clustering of samples (such as vessels in a reference fleet) can have impacts on estimates of age composition (Aanes and Pennington, 2003; ICES, 2008), whilst haul- or trip-based sampling has impacts on accuracy of estimated size distributions (Plet-Hansen *et al.*, 2020). Such biases can be identified by comparing sampling with another data source of known reliability (e.g. Starr and Vignaux, 1997).

Improving future estimations

For estimating total catches of all species in a fishery, an overly simple evaluation of bias is not appropriate. How much bias can be tolerated is dependent on a wide range of factors including intended use, accompanying precision, and sampling limitations, all of which vary between species.

Our study aimed to identify potential biases when estimating total catches. First and foremost, vessel-related biases were present in estimates using probabilistic sampling of vessels, suggesting the importance of accounting for vessel clustering of samples. Our results also suggest that the ratio estimator may be biased across all species. The variations between species are very complex and cannot be explained by encounter rates or commercial importance. There are many other species-related factors that affect bias in estimations of total catches such as patterns in spatial distribution. Therefore, alternative estimators should also be considered for individual species where assumptions of the ratio estimator are violated (Lohr, 2010), such as with rare species where the delta lognormal estimator may be more appropriate (Pennington, 1983; Ortiz *et al.*, 2000).

Although estimator bias was not directly accounted for in the random forest models, the degree to which it affects the overall bias of estimates is evident in the model interpretation (Figure 4). First, estimator bias of simple estimators resulted in an overall small yet detectable positive bias. Particularly large overestimations of reported catches for some rare species, regardless of vessel sampling method (e.g. lesser redfish in the longline fishery; Figure 4c) show that the severity of estimator bias is complex and may impact some species more than others. Fortunately, estimator bias can be evaluated in a future study in the direct context of total catches to fully understand the severity of ignoring the hierarchical sampling design.

In addition to determining the most appropriate design-based estimator to reduce bias, an improvement in sampling design may also reduce bias in estimates of total catches. Our results show that whilst the expert judgement selection of Norwegian Reference Fleet vessels behaves like a simple random sample for many species (Figure 4c), this assumption does not hold for all species. Increasing incentives for vessels to apply to open tenders would improve the expert judgement selection to identify the most "typical" vessels. Furthermore, in the event of multiple eligible vessels, then a larger pool of vessels would be available for the final random selection.

This study focused on the accuracy of estimators, but it is also important to consider the precision. Assumptions behind design-based estimators are different for defining sampling variance formulae (Lohr, 2010) and resampling methods for bootstrapping of variance estimates. The vessel selection biases identified in this study highlight the limitations of a small, fixed sample of vessels. For the limited number of species where bias could be deemed excessive, it is arguably better to relax assumptions in variance estimation to give a more conservative estimate in the aim of being vaguely right rather than precisely wrong.

Data availability statement

The data underlying this article were provided by the Norwegian Directorate of Fisheries by permission. Data will be shared on request to the corresponding author with permission of the Norwegian Directorate of Fisheries.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Funding

This research was funded equally by the Norwegian Institute of Marine Research and the Norwegian Directorate of Fisheries.

Acknowledgements

We thank Hans Hagen Stockhausen and Geir Blom for their helpful comments on draft versions of the manuscript.

References

Aanes, S., and Pennington, M. 2003. On estimating the age composition of the commercial catch of northeast Arctic cod from a sample of clusters. ICES Journal of Marine Science, 60: 297–303.

Ainsworth, C. H., and Pitcher, T. J. 2005. Estimating illegal, unreported and unregulated catch in British Columbia's marine fisheries. Fisheries Research, 75: 40–55.

Bærum, K. M., Anker-Nilssen, T., Christensen-Dalsgaard, S., Fangel, K., Williams, T., and Vølstad, J. H. 2019. Spatial and temporal variations in seabird bycatch: incidental bycatch in the Norwegian coastal gillnet-fishery. Plos ONE, 14: e0212786.

Bell, R. J., Gervelis, B., Chamberlain, G., and Hoey, J. 2017. Discard estimates from self-reported catch data: an example from the U.S. northeast shelf. North American Journal of Fisheries Management, 37: 1130–1144.

Benoît, H. P., and Allard, J. 2009. Can the data from at-sea observer surveys be used to make general inferences about catch composition and discards? Canadian Journal of Fisheries and Aquatic Sciences, 66: 2025–2039.

Berg, H. S. F., and Nedreaas, K. 2020. Estimering av utkast i Norsk kystfiske med garn 2012-2018. Institute of Marine Research.

Borges, L., Zuur, A. F., Rogan, E., and Officer, R. 2004. Optimum sampling levels in discard sampling programs. Canadian Journal of Fisheries and Aquatic Sciences, 61: 1918–1928.

Boulesteix, A.-L., Janitza, S., and Kruppa, J. 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. WIREs Data Mining and Knowledge Discovery 2: 493–507.10.1002/widm.1072.

Bowering, R., Storr-Paulsen, M., Tingley, G., Bjørkan, M., Vølstad, J. H., Gullestad, P., and Lorentsen, E. 2011. Evaluation of the Norwegian Reference Fleet. Institute of Marine Research, Bergen. http://hdl.handle.net/11250/116704 (last accessed 10 June 2021).

Breiman, L. 2001. Random forests. Machine Learning, 45: 5–32.

Burnham, K. P., and Anderson, D. R. 2002. Model selection and multimodel inference: A practical information-theoretic approach. Springer, New York, NY. 515pp.

Cahalan, J., and Faunce, C. 2020. Development and implementation of a fully randomized sampling design for a fishery monitoring program. Fishery Bulletin, 118: 87–99.

Cahalan, J. A., Gasper, J., Mondragon, J., and Trenkel, V. 2015. Catch estimation in the federal trawl fisheries off Alaska: a simulation approach to compare the statistical properties of three trip-specific catch estimators. Canadian Journal of Fisheries and Aquatic Sciences, 72: 1024–1036.

Clegg, T., and Williams, T. 2020. Monitoring bycatches in Norwegian fisheries - Species registered by the Norwegian Reference Fleet. Rapport fra havforskningen; 2020-8. https://hdl.handle.net/11250/2685855 (last accessed 10 June 2021).

Cosandey-Godin, A., Krainski, E. T., Worm, B., and Flemming, J. M. 2015. Applying Bayesian spatiotemporal models to fisheries bycatch in the Canadian Arctic. Canadian Journal of Fisheries and Aquatic Sciences, 72: 186–197.

Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. J. 2007. Random forests for classification in ecology. Ecology, 88: 2783–2792.

Diamond, S. L. 2003. Estimation of bycatch in shrimp trawl fisheries: a comparison of estimation methods using field data and simulated data. Fishery Bulletin, 101: 484–500.

Ewell, C., Hocevar, J., Mitchell, E., Snowden, S., and Jacquet, J. 2020. An evaluation of regional fisheries management organization at-sea compliance monitoring and observer programs. Marine Policy, 115: 103842.

Fangel, K., Aas, Ø., Vølstad, J. H., Bærum, K. M., Christensen-Dalsgaard, S., Nedreaas, K., Overvik, M. *et al.* 2015. Assessing incidental bycatch of seabirds in Norwegian coastal commercial fisheries: empirical and methodological lessons. Global Ecology and Conservation, 4: 127–136.

Faunce, C. H. 2011. A comparison between industry and observer catch compositions within the Gulf of Alaska rockfish fishery. ICES Journal of Marine Science, 68: 1769–1777.

Fernandes, A. C., Oroszlányová, M., Silva, C., Azevedo, M., and Coelho, R. 2021. Investigating the representativeness of onboard sampling trips and estimation of discards based on clustering. Fisheries Research, 234: 105778.

Fox, E. W., Hill, R. A., Leibowitz, S. G., Olsen, A. R., Thornbrugh, D. J., and Weber, M. H. 2017. Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. Environmental Monitoring and Assessment, 189: 316.

Friedman, B. J. H. 2001. Greedy function approximation: a graident boosting machine. The Annals of Statistics, 29: 1189–1232.

Genuer, R., Poggi, J.-M., and Tuleau-malot, C. 2009. Variable selection using random forests. Pattern Recognition Letters, 117: 256–266.

Gezelius, S. S. 2006. Monitoring fishing mortality: compliance in Norwegian offshore fisheries. Marine Policy, 30: 462–469.

Gullestad, P., Blom, G., Bakke, G., and Bogstad, B. 2015. The "Discard ban package": experiences in efforts to improve the exploitation patterns in Norwegian fisheries. Marine Policy, 54: 1–9.

Hothorn, T., Hornik, K., and Zeileis, A. 2006. Unbiased recursive partitioning: a conditional inference framework. Journal of Computational and Graphical Statistics, 15: 651–674.

ICES. 2008. Report of the Workshop on Methods to Evaluate and Estimate the Accuracy of Fisheries Data used for Assessment (WKACCU). 27–30 October 2008 Bergen, Norway. ICES CM 2008\ACOM:32. REF. LRC, RMC. PGCCDBS.

Jacobsen, R. B., Wilson, D. C. K., and Ramirez-Monsalve, P. 2012. Empowerment and regulation - dilemmas in participatory fisheries science. Fish and Fisheries, 13: 291–302.

Jessen, R.J. 1978. Statistical Survey Techniques. John Wiley & Sons, New York, NY.

Jones, Z. M., and Linder, F. J. 2016. edarf: exploratory data analysis using random forests. The Journal of Open Source Software, 1: 92.

Kraan, M., Uhlmann, S., Steenbergen, J., van Helmond, A. T. M., and van Hoof, L. 2013. The optimal process of self-sampling in fisheries: lessons learned in the netherlands. Journal of Fish Biology, 83: 963–973.

Liggins, G. W., Bradley, M. J., and Kennelly, S. J. 1997. Detection of bias in observer-based estimates of retained and discarded catches from a multi species trawl fishery. Fisheries Research, 32: 133–147.

Lohr, S. L. 2010. Sampling: Design and Analysis. Chapman and Hall/CRC. 596pp.

Lordan, C., Ó Cuaig, M., Graham, N., and Rihan, D. 2011. The ups and downs of working with industry to collect fishery-dependent data: the irish experience. ICES Journal of Marine Science, 68: 1670–1678.

Mangi, S. C., Dolder, P. J., Catchpole, T. L., Rodmell, D., and de Rozarieux, N. 2015. Approaches to fully documented fisheries: practical issues and stakeholder perceptions. Fish and Fisheries, 16: 426–452.

Mangi, S. C., Smith, S., and Catchpole, T. L. 2016. Assessing the capability and willingness of skippers towards fishing industry-led data collection. Ocean and Coastal Management, 134: 11–19.

Moan, A., Skern-mauritzen, M., Vølstad, J. H., and Bjørge, A. 2020. Assessing the impact of fisheries-related mortality of harbour porpoise (*Phocoena phocoena*) caused by incidental bycatch in the dynamic Norwegian gillnet fisheries. ICES Journal of Marine Science, 77: 3039–3049.

Nelson, A. G. 2014. Cluster sampling: a pervasive, yet little recognized survey design in fisheries research. Transactions of the American Fisheries Society, 143: 926–938.

Norwegian Directorate of Fisheries. 2021. Norske Omregningsfaktorer - For omregning av landet produktvekt av marin fisk og andre marine arter til rund vekt - for fiske i det nordlige Atlanterhavet og i andre farvann. https://www.fiskeridir.no/Yrkesfiske/Tema/Omregningsfaktorer (last accessed 10 June 2021).

Ortiz, M., Laegault, C. M., and Ehrhardt, N. M. 2000. An alternative method for estimating bycatch from the U.S. shrimp trawl fi shery in the Gulf of Mexico, 1972–1995. Fishery Bulletin, 98: 583–599.

Pennington, M. 1983. Efficient estimators of abundance for fish and plankton surveys. Biometrics, 39: 281–286.

Pitcher, T. J., Watson, R., Forrest, R., Þór Valtýsson, H., and Guénette, S. 2002. Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish and Fisheries, 3: 317–339.

Plet-Hansen, K. S., Bastardie, F., and Ulrich, C. 2020. The value of commercial fish size distribution recorded at haul by haul compared to trip by trip. ICES Journal of Marine Science, 77: 2729–2740.

Rochet, M.-J., and Trenkel, V. M. 2005. Factors for the variability of discards: assumptions and field evidence. Canadian Journal of Fisheries and Aquatic Sciences, 62: 224–235.

Roman, S., Jacobson, N., and Cadrin, S. X. 2011. Assessing the reliability of fisher self-sampling programs. North American Journal of Fisheries Management, 31: 165–175.

Siders, Z., Ducharme-Barth, N., Carvalho, F., Kobayashi, D., Martin, S., Raynor, J., Jones, T. *et al.* 2020. Ensemble random forests as a tool for modeling rare occurrences. Endangered Species Research, 43: 183–197.

Snyder, H. T., and Erbaugh, J. T. 2020. Fishery observers address arctic fishery discards. Environmental Research Letters, 15: 0940c4.

Starr, P. 2010. Fisher-Collected sampling data: lessons from the New Zealand experience. Marine and Coastal Fisheries, 2: 47–59.

Starr, P. J., and Vignaux, M. 1997. Comparison of data from voluntary logbook and research catch-sampling programmes in the New Zealand lobster fishery. Marine and Freshwater Research, 48: 1075–1080.

Stock, B. C., Ward, E. J., Eguchi, T., Jannot, J. E., Thorson, J. T., Feist, B. E., and Semmens, B. X. 2020. Comparing predictions of fisheries bycatch using multiple spatiotemporal species distribution model frameworks. Canadian Journal of Fisheries and Aquatic Sciences, 77: 146–163.

Stratoudakis, Y., Fryer, R. J., Cook, R. M., and Pierce, G. J. 1999. Fish discarded from Scottish demersal vessels: estimators of total discards and annual estimates for targeted gadoids. ICES Journal of Marine Science, 56: 592–605.

Strobl, C., Boulesteix, A. L., Zeileis, A., and Hothorn, T. 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics, 8: 25.

Strobl, C., Boulesteix, A., Kneib, T., Augustin, T., and Zeileis, A. 2008. Conditional variable importance for random forests. BMC Bioinformatics, 11: 307.

Strobl, C., Hothorn, T., and Zeileis, A. 2009. Party on! A new, conditional variable importance measure available in the party package. The R Journal, 1: 14–17.

Suuronen, P., and Gilman, E. 2020. Monitoring and managing fisheries discards: new technologies and approaches. Marine Policy, 116: 103554.

Tamsett, D., Janacek, G., Emberton, M., Lart, B., and Course, G. 1999. Onboard sampling for measuring discards in commercial fishing based on multilevel modelling of measurements in the irish sea from NW England and N Wales. Fisheries Research, 42: 117–126.

Vølstad, J. H., and Fogarty, M. 2006. Report on the National Observer Program vessel selection bias workshop Woods Hole, May 17-19, 2006.

Handling Editor: Ernesto Jardim

PAPER III

Title

Evaluating assumptions behind design-based estimators for unreported catches

Authors

Thomas L Clegg*, Edvin Fuglebakk, Kotaro Ono

Institute of Marine Research/Havforskningsinstituttet, Postboks 1870 Nordnes, NO-5817 Bergen, Norway

* Corresponding author (tom.clegg@hi.no)

Abstract

Understanding a fishery's impact on the marine ecosystem requires a quantification of total catches, which include unreported catches. For recent years in Norwegian waters, unreported catches have been estimated using data collected by the Norwegian Reference Fleet, a fisher self-sampling programme that regularly gathers data on catches of all species. In this study, we focused on the use of design-based estimators for total catches in offshore fisheries, which have previously been used to estimate discards in the Norwegian coastal gillnet fisheries. After adapting the current methodology to the data available in offshore fisheries, we explored the assumptions behind both unit- and ratio-based estimators, and the effect of ignoring the clustered nature of data. Using a jack-knife resampling method to estimate the true bias in estimates of total catches and associated variability, we found that ignoring the clustered nature of data tended to underestimate the variability, which lead to occurrences where unreported catches were statistically detected when in fact there was too much uncertainty to make such a conclusion. Further validations suggested the assumptions of the cluster unit estimator are likely not met due to a non-random vessel selection that favours more active vessels. We therefore concluded that the unit and ratio cluster estimator is applied and compared, as per best practice routines.

Keywords

unreported catches; design-based estimators; reference fleet; cluster; self-sampling

1. Introduction

Unwanted catches are an inevitable consequence of selective fishing. Supportive policies can increase avoidance or utilisation of unwanted catches (Karp et al., 2019), but no fishing gear is perfectly selective and discarding of low value catches is incentivised through targeting of highly valuable species (Batsleer et al., 2015). A discard ban is a central part of fisheries policy aimed at reducing unwanted catches, which is typically accompanied by a landing obligation. Ensuring that all catches are landed and reported improves an understanding of the environmental impacts of fisheries. However, even with relatively high compliance rates (Gezelius, 2006; Hønneland, 2000), there are still sources of unreported catches (Box 1) that should be accounted for.

Box 1. Definitions

Total catch	The biomass of marine resources that are brought on board the vessel.
Landings or retained catch	The retained portion of total catches that is landed and reported through mandatory channels.
Bycatch	The portion of catch of non-target species, which can either be landed or discarded.
Discards	That portion of the total catch which is thrown away or dumped at sea before landing for whatever reason. It includes incidental catches such as marine mammals and seabirds, but does not include shells, corals, plants, or inorganic materials, nor processing waste such as offal and carcasses.
Unreported catches	The portion of total catches that are not explicitly reported to species level in official statistics. Unreported catches comprise of discards, illegal catches, and unmandated catches (Pitcher et al., 2002)

Since 1987 when a discard ban was introduced in Norway, a broad range of technical measures have been subsequently implemented to improve selectivity and encourage full utilisation of fisheries (Gullestad et al., 2014), collectively referred to as the 'Discard Ban Package' (see Gullestad *et al.*, 2015 for full description). Snapshot studies and historical reconstructions of unreported catches have found that discards have decreased since the introduction of the discard ban (Dingsør, 2001; McBride and Fotland, 1996; Nedreaas et al., 2015). This decrease is supported by indirect evidence that the Discard Ban Package has improved the status of many commercially important fish stocks (Diamond and Beukers-Stewart, 2011; Gullestad et al., 2014). It is therefore typically assumed that misreported catches are negligible in Norwegian fisheries (Gilman et al., 2020; ICES, 2021; Kelleher, 2005; Pérez Roda et al., 2019) despite the acknowledgement that discarding still occurs in Norwegian waters, and that current levels of misreported catches are mostly unknown (Gezelius, 2006; Gullestad et al., 2015; Nedreaas et al., 2015).

In recent years, unreported catches have been estimated in Norway using data collected by the Norwegian Reference Fleet, a group of active fishing vessels trained and paid to self-sample their catches (Clegg and Williams, 2020). Coastal vessels participating in the Norwegian Reference Fleet record discards explicitly, allowing for estimates of bycatches and discards of all species in coastal gillnet fisheries (fish: Berg and Nedreaas, 2020; seabirds: Fangel *et al.*, 2015; Bærum *et al.*, 2019; marine mammals: Moan, 2016; Moan *et al.*, 2020). These studies have used a combination of design- and model-based approaches to understand the scale and drivers behind discarding. However, current routines have not yet been adapted to offshore fisheries where the Norwegian

Reference Fleet do not explicitly record discards, but instead report total catches (i.e., landed and unreported catches combined).

The Norwegian Reference Fleet programme has a complex and clustered sampling design to account for the voluntary, long-term participation of vessels, multiple sampling objectives, and practical constraints of fisheries sampling. Of specific importance to this study is the clustering of samples within vessels. Ignoring the clustered nature of sampling designs can have large impacts on both the bias and precision of estimates (Lohr, 2010; Nelson, 2014), increasing the risk of concluding that unreported catches are significantly high. The representativeness of a non-random vessel selection is unknown statistically, but can be inferred through theoretical sampling design or by evaluating estimator assumptions.

Here we present a revised estimator for unreported catches in Norwegian fisheries. We evaluated the assumptions behind the estimator currently used for unreported catches in Norwegian fisheries (Bærum et al., 2019; Berg and Nedreaas, 2020; Fangel et al., 2015), and demonstrate why accounting for clustered nature of sampling by the Norwegian Reference Fleet should not be ignored. We then evaluate the representativeness of data with regards to estimators based on cluster sampling, which identified that a ratio estimator is more reliable given the potential biases in non-random sampling of vessels. For this study, we used a historical case study of the Barents Sea longline fishery. We therefore discuss how applicable the proposed estimator is to other fisheries in which the Norwegian Reference Fleet is active and identify aspects of the estimator that still need development to ensure a reliable estimate of unreported catches.

2. Material and methods

2.1. Case study fishery and species

The Barents Sea longline fishery is defined in this study as vessels over 28 m overall length operating in the statistical areas shown in Figure 1. The fishery operates over almost the entire Barents Sea, but we restrict our study to the statistical areas where vessels are most active. The fishery operates year-round but is restricted in northern areas in winter months by expanding sea ice cover. Vessels predominantly target cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*), but also target a wide range of other demersal fish species, notably tusk (*Brosme brosme*), Greenland halibut (*Reinhardtius hippoglossoides*), and wolffish (*Anarhichas* spp.).

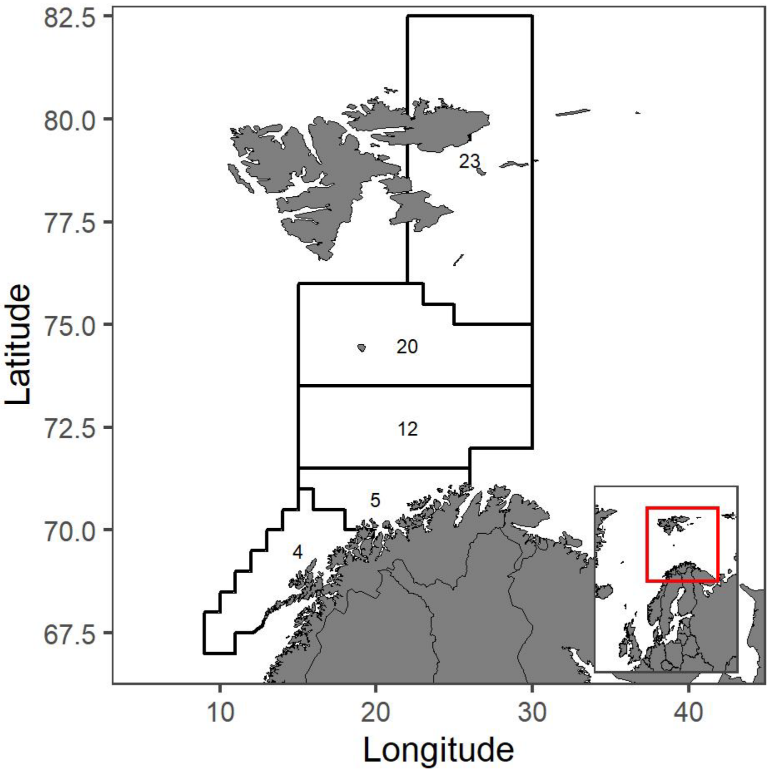


Figure 1. Statistical areas in the Barents Sea longline fishery as defined by the Norwegian Directorate of Fisheries.

This study focused on fish species, therefore excluding birds, sea mammals, and invertebrates. Extremely large catches were removed if there was no verification, using expert knowledge and calculating z-scores for each species to manually identify outliers.

A total of 50 species or species groups were observed by the Norwegian Reference Fleet in the Barents Sea longline fishery (See Supplementary Materials for full list). For some species groups, mandatory catch reporting does not require species-level identification. On the other hand, the Norwegian Reference Fleet samplers are trained to identify all individuals to species level. For example, in the study fishery, only 1.5 % of landed skates and rays (order: Rajiformes) by weight were reported to species level between 2012 and 2018. Comparatively, the Norwegian Reference Fleet identified 98.2% of skates and rays to species level. We therefore removed skate and ray observations from 121 sampled fishing days where a species was not identified. This removal reduces the sample size for skate and ray species but allows us to estimate unreported catches for individual species. We assumed that all unidentified redfish species (genus: *Sebastes*; 0.5 % of total sampled weight) were the lesser redfish (*Sebastes viviparus*). Where species groups are reported in official catch statistics, we estimated total catches for individual species in the group, which can then be aggregated to the desired taxonomic level for comparison with reported catches.

2.2. Data

2.2.1. Norwegian Reference Fleet

Vessel selection is required by law to follow a public tender process. Each tender lists required and desired specifications that a vessel must meet to be eligible for participation. An expert panel reviews each application, evaluating whether the vessel meets the required and desired criteria based on evidence of fishing activity provided by the applicant. If there are multiple eligible vessels after this review process, the contract is awarded randomly. The criteria listed in tenders and expert judgement selection process aims to simulate a stratified random sample of vessels from the fishery.

Vessels sample total catches for one fishing operation every two days, known as a 1-in-2 systematic sample. A systematic sample is expected to behave like a simple random sample (Lohr, 2010), and has been proven for the Norwegian Reference Fleet in the context of reported catches (Clegg et al., 2022). On each sampling day, total catches are recorded from three representative samples of consecutive hooks are taken from the start, middle, and end of longlines, that the crew or skipper deems representative of the catch composition for that day. Therefore, total catches per day are extrapolated using the total:sampled hook ratio.

We used 5 484 observed fishing days from six vessels between 2012 and 2018, which equates to almost 10 % of all fishing days in the study period from 16 % of vessels in the fishery (Table 1). Total hooks could not be determined for 67 sampled fishing days (3 %), due to either erroneous misreporting or sampling over midnight such that dates did not match. We imputed these values with the modal number of hooks used by that vessel in the study period.

Table 1. Summary of sampling by the Norwegian Reference Fleet sampling programme in the Barents Sea longline fishery. The summary across all years is not the sum of individual years because the same vessels are active over multiple years.

Year	Vessels			Fishing days		
	Sample	Population	Sampling fraction	Sample	Population	Sampling fraction
2012	6	36	0.17	758	4 943	0.150
2013	5	34	0.15	320	3 471	0.092
2014	6	27	0.22	224	2 998	0.075
2015	4	27	0.15	176	2 698	0.065
2016	4	28	0.14	206	3 172	0.065
2017	4	26	0.15	158	2 866	0.055
2018	4	28	0.14	274	2 952	0.093
All years	6	42	0.14	2 116	23 100	0.092

2.2.2. Daily logbooks (Electronic Reporting System)

In the Barents Sea, Norwegian fishing vessels longer than 15 m overall length are required to keep an up-to-date logbook of catches and fishing activity using an electronic reporting system. A catch report must be sent at least once per calendar day and is required for each fishing operation (defined as the period from the fishing gear entering the water until it is taken out of the water). However, for passive gears such as longlines and gillnets where it is more difficult to define discrete fishing operations, catch reports are typically sent as a single daily summary.

A description of fishing effort is included for each fishing operation, which depends on the fishing gear used. To generalise, fishing duration of active gears is calculated as the difference between the start and stop time of the fishing operation. For passive gears, the total number of nets or hooks is reported per calendar day. Alongside fishing activity, skippers are also required to maintain an up-to-date estimate of catches from each fishing operation. However, as these logbooks are used for control and enforcement of the fishery, they only contain the retained portion of catches. The accuracy of logbook information is maintained through inspections of storage holds at sea and catches officially reported upon landing (see *Sales notes* below). Weights are estimated on-board but must be within 10 % of the official weight reported upon landing. Furthermore, species reporting at sea is not as strict as upon landing, meaning many species are often grouped. Due to these uncertainties in reported catch estimates, we concluded that logbook catches are not useful for comparison with estimated total catches to infer unreported catches.

Daily logbooks were provided by the Norwegian Directorate of Fisheries. We used them as a measure of fishing effort for all fishing days by vessels in the Barents Sea longline fishery, which were provided by the Norwegian Directorate of Fisheries. Of the 23 100 logbook entries, 231 (1 %) reported less than 1 000 hooks used. We deemed these entries erroneous and therefore imputed the number of hooks used as the modal value for that vessel in the study period.

2.2.3. Sales notes

All first-hand sales of fish are directed through one of six sales organisations in Norway (reduced to five in 2020). Upon landing, a sales note must be immediately sent to the sales organisation to receive payment for catches. These sales notes are also sent to the Norwegian Directorate of Fisheries as the official record of catches, who provided them for this study. Reported weights are recorded using officially approved scales, and the sales note is signed by both buyer and seller to reduce the opportunity for fraudulent reporting. The sales organisations are responsible for confiscating illegal catches, monitoring quota, and reporting vessels in breach of fisheries law. Sales organisations are subject to on-site or data inspection at any time. This centralised system provides the most reliable data source on reported catches, which we deducted from estimated total catches to infer unreported catches.

Reported weights of fish are recorded after any processing on board, and therefore require conversion back into the round weight (live weight when removed from the water). Conversion factors are intermittently published as annual mean values for all areas (Norwegian Directorate of Fisheries, 2021). Due to the difficulties in quantifying uncertainties in conversion factors for all species and products, we assume reported round weight landings are without error.

2.3. Standard estimation framework

The estimation routine started with post-stratifying samples into strata defined as a combination of year, statistical area (Figure 1) and season (winter: January-April; summer: May-August; autumn: September-December). The defined estimator (Table 2) was then applied to each stratum individually. All estimators were applied to total catches (i.e., before sorting), as offshore vessels in the Norwegian Reference Fleet only began reporting discarded and retained portions of the catch in 2019. Unreported catches must therefore be inferred by deducting catches reported in sales notes.

Table 2. Candidate estimators for unreported catches using Norwegian Reference Fleet data. Estimators were applied to individual strata, defined as year, statistical area (Figure 1) and annual quarter. See Table 3 for notation in formulae.

Estimator	Equation		Assumptions
Simple			
Unit	$\hat{Y} = \frac{M}{m} \sum_{j=1}^{M} y_j$	(1)	• Primary sampling unit = fishing day • Observations are a simple random sample of all fishing days

	Cluster	$\hat{Y} = X\frac{\sum_{j=1}^{m} y_j}{\sum_{j=1}^{m} x_j}$	(2)	• Ratio: strong correlation between total catches (y_j) and fishing effort (x_j) for individual fishing operations
Cluster				
	Unit	$\hat{y}_i = \frac{M_i}{m_i}\sum_{j=1}^{m_i} y_{ij}$	(3.1)	• Primary sampling unit = vessel • Secondary sampling unit = fishing day • Observed vessels are a simple random sample from all vessels. • Observed fishing days are a simple random sample from each vessel • Ratio: strong correlation between total catches (y_i) and fishing effort (x_i) for individual vessels
		$\hat{Y} = \frac{N}{n}\sum_{i=1}^{n} \hat{y}_i$	(3.2)	
	Ratio	$\hat{y}_i = \frac{X_i}{x_i}\sum_{j=1}^{m_i} y_{ij}$	(4.1)	
		$\hat{Y} = X\frac{\sum_{i=1}^{n} \hat{y}_i}{\sum_{i=1}^{n} X_i}$	(4.2)	

Table 3. Notation for Equations in Table 2.

Notation	Sample	Population
Weight of total catches*	y	Y
Fishing effort	x	X
Number of vessels	n	N
Reference to vessel	$i\ (i = 1, \dots, n)$	$i\ (i = 1, \dots, N)$
Number of fishing operations	m	M
Reference to fishing operation	$j\ (j = 1, \dots, m)$	$j\ (j = 1, \dots, M)$
Reference to jack-knife replicate	$k\ (k = 1, \dots, K)$	
Reference to bootstrap replicate	$b\ (b = 1, \dots, B)$	

* Estimated sample and population totals denoted as $\hat{y}$ and $\hat{Y}$ respectively

The current methodology for estimating discards and bycatches in Norwegian fisheries is based on simple estimators (Equations 1 and 2; Table 2). These estimators assume that fishing operations were a simple random sample from all fishing activity on the level of each stratum. Furthermore, the ratio estimator assumes a strong correlation between total catches and fishing effort for individual fishing operations. The simple ratio estimator is an extension of the simple unit estimator which assumes a relationship between total catches and fishing effort for individual fishing operations. The ratio estimator is expected to improve precision at the expense of some expected bias (Lohr, 2010). In the longline fishery, we defined fishing effort as the number of hooks used per calendar day.

We defined two additional estimators based on cluster sampling, which better reflects the sampling design of the Norwegian Reference Fleet. Cluster estimators identify vessels as the primary sampling unit, which are assumed to be a simple random selection from the fishing fleet. Fishing days are secondary sampling units, which are assumed to be a simple random selection from all fishing days by each vessel individually. For the cluster ratio estimators, fishing effort was again defined as number of hooks used.

We defined a stratum as unsampled if it had less than three observed fishing days. Total catch rates in unsampled strata were imputed by borrowing data from adjacent strata which are assumed to have similar rates. We defined a three-tier imputation routine where for each unsampled stratum, we gradually expanded the strata from which we borrowed data until we had sufficient data for an estimation. Firstly, we borrowed data from the same statistical area and season in the years adjacent (for 2012 and 2018, the first and last year in the study, we borrowed data from the only adjacent year for which we have data). If there were no observations in adjacent years, we expanded the imputation to include observations from all years in that statistical area and season. If a statistical area was not observed for a given quarter in any years, then we estimated the total catch rate using all data in the study.

Variability of estimated total catches for species were estimated using the bootstrap method (B = 5 000 replicates) (Efron and Tibshirani, 1994). To estimate the variability of simple estimators, we defined a 'simple' bootstrap routine that reflects the estimator assumptions. For each stratum, fishing days were resampled with replacement from a single pool including all vessels, with a sample size equal to the original dataset. The variability of cluster estimators was defined by accounting for potential variation in each sampling stage. For each year, we first resampled Norwegian Reference Fleet vessels with replacement, then post-stratified samples. Then for each stratum, we resampled fishing days with replacement for each vessel individually. We used bootstrap replicates to calculate 95 % confidence intervals using the percentile method. Unreported catches were inferred as the difference between estimated total catches and those officially reported in sales notes. If reported catches fell outside the confidence interval of estimated total catches, we considered unreported catches to be statistically detected.

2.4. Quantifying estimator bias

A typical validation of estimator performance (bias and precision) involves identifying a domain with a true value with which to compare estimates. However, this is not possible for fishery-level estimates of unreported catches. Only observations of total catches are available, from which the unreported portion must be inferred. Even considering observed total catches, a bias assessment is complicated by the sub-sampling of hooks which means we are not even certain of total catches for any sampling unit (vessels or fishing days). Finally, for strata in which only one vessel was sampled, removing that vessel would result in no observations for an estimation.

The domain for testing biases was defined as total annual catches by Norwegian Reference Fleet vessels in strata in which two or more vessels were sampled (Y^*). This testing domain involves firstly extrapolating sampled hooks to the day-level, and secondly to the vessel-level for each stratum before being summed. This first extrapolation step is a typical necessity where sampling of large fishing operations is unfeasible. The second step is necessary such that the 'truth' is defined at the level of primary sampling units (vessels), which then allows for resampling of secondary sampling units (fishing days). Sub-sampling of fishing operations is common and extrapolations are often done on-board to estimate the haul-level catches (e.g., Borges et al., 2005). The extrapolation to vessel-level catches per strata is also assumed to introduce negligible bias due to the robust systematic sampling routines for fishing operations for each vessel (Clegg et al., 2022). Limiting strata to those with two or more sampled vessels avoids imputation of under-sampled strata, which will introduce additional biases. Given that this testing is limited to sampling data, the evaluation focuses on estimator biases by excluding selection biases which relate to the representativeness of samples.

The jack-knife resampling method was defined as followed (see Figure 2 for schematic diagram): for each year, a single vessel was randomly removed from the dataset. Then, fishing days were resampled randomly with replacement for each vessel and stratum. This resampled dataset (k) was then used to re-estimate total sampled catches ($\hat{Y}^*_k$) for each species using the estimators defined in Table 2. The jack-knife resampling process was repeated K = 5 000 times, which was sufficient to approximate an equal number of removals for all vessels. Total catches were also estimated for the testing domain dataset using the bootstrap method ($\hat{Y}^*_b$; see standard estimation framework section), using B = 5 000 bootstrap replicates.

Biases in the design-based estimators (Table 2) and associated variability were calculated by comparing the jack-knife estimates with the truths we have defined as followed:

- Bias in each design-based estimator was calculated using Equation 5, defined as the mean relative error of jack-knife estimates (MRE_{est}; Figure 2).
- Bias in the variability of each design-based estimator was calculated using Equation 8, which compares the estimated coefficient of variation (Equation 6) with the true relative error (Equation 7) which we assumed using the jack-knife method.

The cluster unit estimator is theoretically unbiased given random sampling (Lohr, 2010). However, due to the low number of vessels in our dataset, we expected some deviation from zero for even an unbiased estimator, as we cannot simulate enough sampling variation to approximate a continuous distribution.

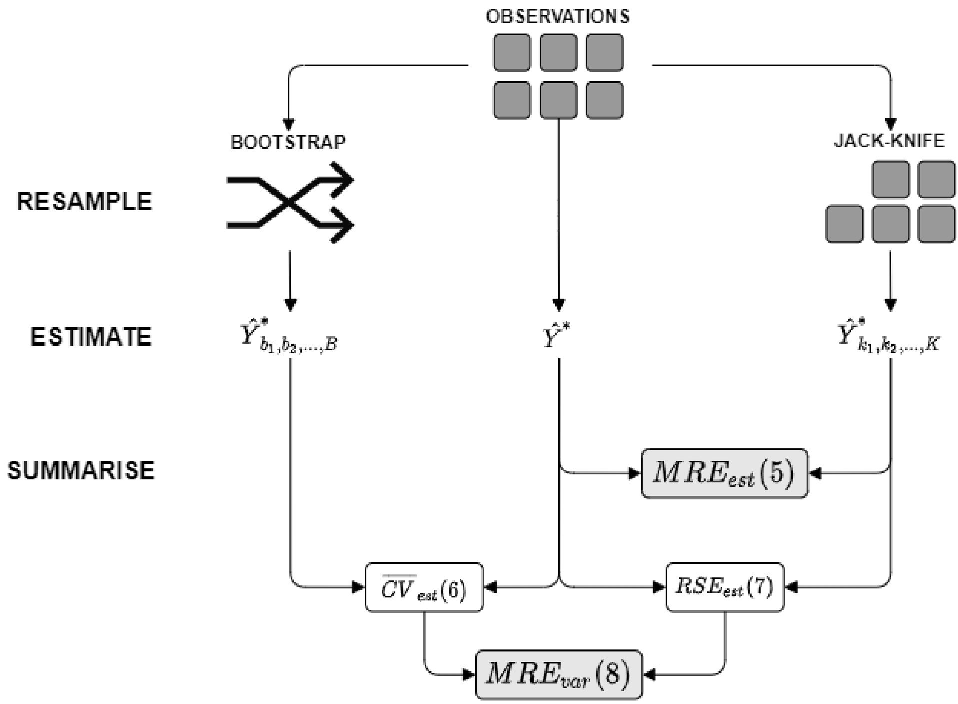


Figure 2. Schematic of bias evaluation. Equations numbers given in brackets.

$$MRE_{est} = \frac{1}{Y^*K}\sum_{k=1}^{K}\left(\hat{Y}^*{}_k - Y^*\right) \qquad (5)$$

$$\overline{CV}_{est} = \frac{1}{Y^*}\sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{Y}^*{}_b - Y^*\right)^2} \qquad (6)$$

$$RSE_{est} = \frac{1}{Y^*}\sqrt{\frac{1}{K-1}\sum_{k=1}^{K}\left(\hat{Y}^*{}_k - Y^*\right)^2} \qquad (7)$$

$$MRE_{var} = \frac{\overline{CV}_{est} - RSE_{est}}{RSE_{\ est}} \qquad (8)$$

The bias in both the estimate (Equation 5) and associated variability (Equation 8) were compared across all four estimators (Table 2) for each species and year to get a generalised overview of estimator performance and more specifically determine the importance of accounting for the clustered nature of data. Based on this analysis, we focused the rest of the analysis on the cluster estimators.

To investigate how biases varied across species, we then plotted the estimated bias (Equation 5) and variance (Equation 6) of the cluster estimators against the encounter rate of species in sampled catches. The sampling design of the Norwegian Reference Fleet is generalised for all species, which will result in varied performance of estimators across species, particularly for rare bycatch species (Martin et al., 2005; Pennington, 1996). Viewing the estimator performance across the range of species encounter rates can help to determine if there is a tolerable limit to estimator performance for rare species.

Finally, we evaluated the core assumption of the cluster ratio estimator: It assumes a linear relationship between total catches and number of hooks. We therefore calculated the Pearson's correlation coefficient (ρ) for each species and year then plotted against the bias (MRE_{est}; Equation 5) and variance (RSE_{est}; Equation 7) of the cluster ratio estimator.

2.5. Assessing the representativeness of samples

The jack-knife resampling method addresses estimator biases but cannot account for selection biases which affect the representativeness of samples. Therefore, we evaluated the representativeness of samples using the best practice method of comparing fishing effort characteristics between samples and the population (ICES, 2007, 2003). We compared total annual fishing days per vessel, for which the cluster unit estimator assumes samples are representative. Unequal fishing days per vessel also indicates that a ratio estimator is more appropriate. We also compared annual mean number of hooks per vessel, which influences the precision of ratio estimators.

2.6. Exploring the chosen estimators

The statistical analyses described above were used to define the best estimators for unreported catches across all species. We explored how the chosen estimation procedure affected management advice for three commercially important species in the Barents Sea longline fishery, namely cod, haddock, and beaked redfish (*Sebastes mentella*), to demonstrate how the sensitivity of statistical detection could influence the ability to make actionable conclusions. Cod and haddock are valuable species for which discards are expected to be negligible (ICES, 2021). Beaked redfish and golden redfish are morphologically similar have partially overlapping habitats, making landing statistics less reliable (ICES, 2021). Beaked redfish is a quota-regulated, whilst golden redfish is only landed using quota set aside for unavoidable bycatches, so the risk of misreporting of these two redfish species are likely interlinked.

3. Results

3.1. Quantifying estimator bias

The jack-knife resampling analysis provides evidence for the importance of accounting for the clustered nature of sampling when using Norwegian Reference Fleet data to estimate total catches. Using observations in sampled strata catches as the testing domain, we found the cluster unit estimator performed best overall with negligible bias across all species, whilst the cluster ratio estimator had relatively similar bias to both simple estimators (Figure 3A). Ignoring clustered sampling resulted in an underestimation of variance for almost all species, which improved when applying cluster estimators, albeit with a small tendency to overestimate variance (Figure 3B).

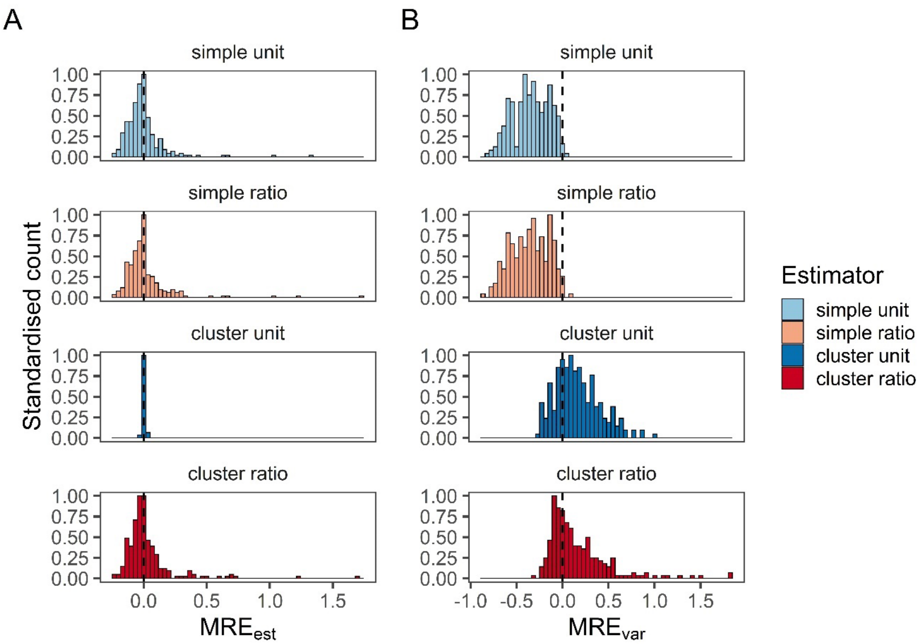


Figure 3. Performance of estimators using total annual observed catches of species by the Norwegian Reference Fleet. (A) Relative error of estimate (Equation 5) and (B) variance (Equation 7). Scaled counts used to compare across estimators.

Whereas the cluster unit estimator is unbiased in all cases, Figure 4A reveals that the cluster ratio estimator is more biased when applied to rarer bycatch species (i.e., low encounter rate). The cluster ratio and unit estimators have similar trends in precision across the range of encounter rates, apart from the rarest species ($\lesssim 10$ % encounter rate) for which the variance is almost twice as large as the mean (RSE_{est} > 2). A poor correlation between total catches and number of hooks begins to affect the performance of the cluster ratio estimator below a threshold of $\rho \approx 0.25$ (, Figure 4B), both with regards to bias and precision.

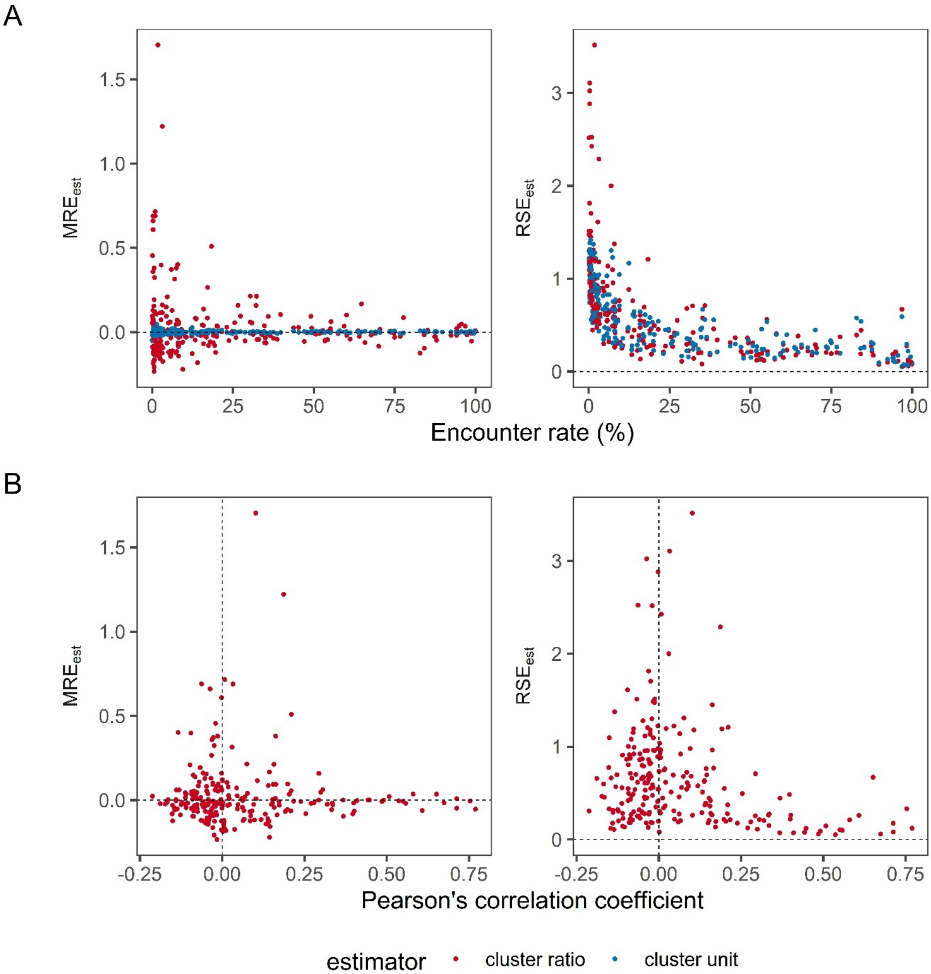


Figure 4. Effect of (A) encounter rate and (B) correlation between total observed catches and fishing effort on the bias (MRE_{est}) and variance (RSE_{est}) of estimators. Each point represents one species in one year. Testing domain limited to observed fishing days in strata where two or more vessels were sampled.

3.2. Assessing the representativeness of samples

Norwegian Reference Fleet vessels are some of the most active vessels in the fishery (Figure 5A), suggesting that samples are not representative of average fishing days per vessel in the fishery. In three of the seven years, the most active vessel has participated in sampling. In addition to a higher number of fishing days, Norwegian Reference Fleet vessels also use more hooks per fishing than most other longline vessels in the fishery (Figure 5B). This combination of using more hooks over more days will lead to likely lead to an overestimation of catches when applying a cluster unit estimator. Comparatively, the cluster ratio estimator accounts for variable fishing effort, meaning that in this regard, it is more tolerant towards the issues in representativeness identified here.

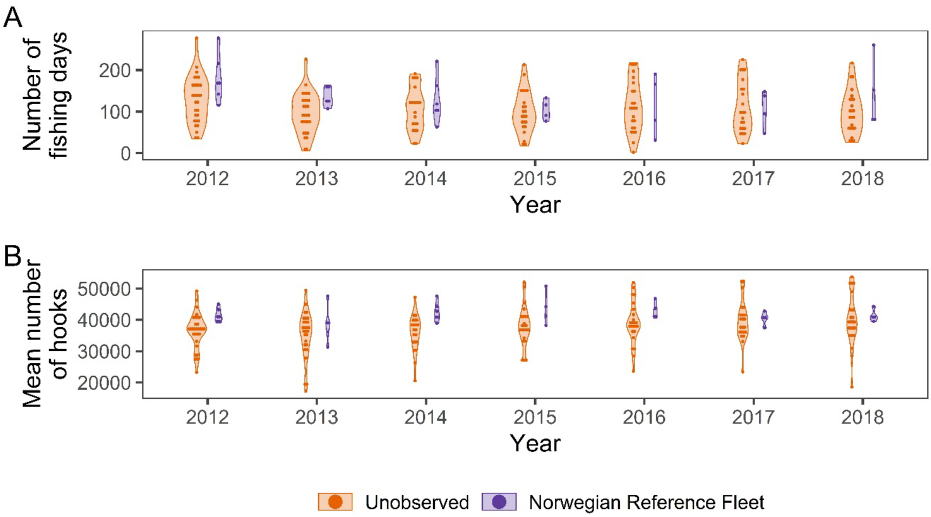


Figure 5. Representativeness of sampled fishing effort in the Barents Sea longline fishery. (A) Number of fishing days and (B) mean number of hooks per fishing operation for each vessel.

3.3. Exploring the chosen estimators

From the evidence presented here, we conclude that the cluster estimators are the best method for estimating unreported catches in Barents Sea longline fishery using data collected by the Norwegian Reference Fleet. The increased uncertainty resulting from accounting for the clustering of data are demonstrated in Figure 6 for selected species. If the simple estimators were applied, then unreported catches of cod would be statistically detected in four of the seven years, and for haddock in five of those years. Conversely, underreporting of golden redfish was statistically detected in 2014 and 2015 if the cluster ratio estimator is applied, compared to the cluster unit estimator for which uncertainty is often larger than the reported component of catches. The tendency for the cluster unit estimator to potentially overestimate total catches is not seen to such a large degree with beaked redfish, indicating that vessel-specific fishing behaviour is highly variable across species.

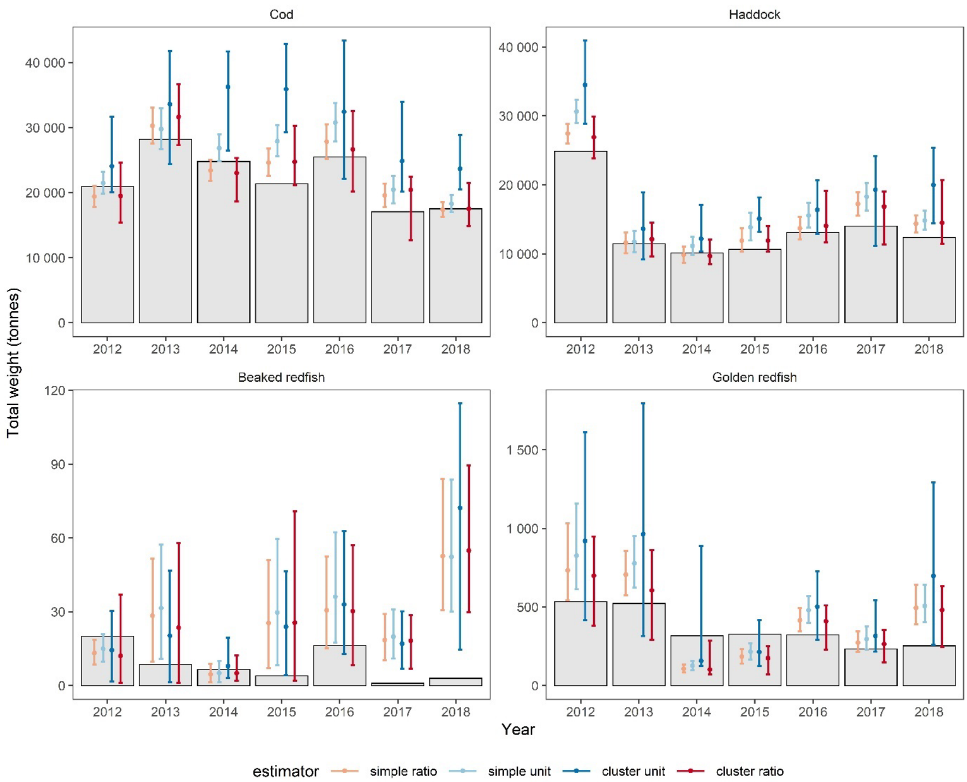


Figure 6. Estimated total annual catches (mean and 95 % confidence interval) of cod (*Gadus morhua*), haddock (*Melanogrammus aeglefinus*), beaked redfish (*Sebastes mentella*), and golden redfish (*Sebastes norvegicus*) compared to reported catches (grey bars), using the four candidate estimators (Table 2).

Deciding between the unit and ratio cluster estimators is not as clear of a conclusion. Whilst the cluster unit estimator is unbiased (Figure 3), the risk of poor representativeness (Figure 5) means that for some species, the cluster unit estimator may have a tendency to overestimate unreported catches relative to all other estimators (Figure 6).

Statistical detection of unreported catches is dependent on the confidence level chosen, which should be considered when interpreting estimates for 50 species. At a 95 % confidence level, there is a 5 % probability that statistically detected unreported catches were a result of chance. Statistical detection of unreported catches is also dependent on the level of aggregation that results are presented. For example, applying the chosen cluster ratio estimator to estimate total unreported catches of haddock in the entire study period results in statistically detectable levels of unreported catches (95 % CI: 1 555–20 734 tonnes), even though for individual years, unreported catches are not statistically detectable (Figure 6).

Final estimates of unreported catches for all species observed in the Barents Sea longline fishery are available in the Supplementary Materials. Total catches of skate and ray species are presented collectively as a species group (order: Rajiformes) to allow for comparison with reported catches, and a separate file presents estimated total catches for individual skate and ray species.

4. Discussion

Using a single estimation routine for unreported catches of all species in a fishery is desirable for the sake of simplicity, speed, and comparability (Gilman et al., 2020; Kennelly, 2020; NMFS, 2011). However, this study has shown that without a sufficient understanding of the bias and uncertainty, the accuracy of estimates across species is unknown and can be highly misleading. The importance of accounting for the clustered nature of fisheries data is well understood (e.g., Aanes and Pennington, 2003; Borges et al., 2005; Fernandes et al., 2021; Lohr, 2010; Nelson, 2014) but Nelson (2014) suggests that clustering is typically ignored due to a lack of awareness. The clustered nature of the Norwegian Reference Fleet sampling design has only been recognised previously when bycatch rates were estimated using a model-based approach (Bærum et al., 2019; Moan et al., 2020) by including vessels as a random intercept in the model, but has been ignored when applying design-based estimators for reasons of simplicity and a historical focus on point estimates rather than on uncertainty. Bias in simple estimators was identified in a previous study in the Barents Sea longline fishery in the context of reported catches (Clegg et al., 2022). Our study supports this finding by demonstrating that cluster-based estimators improve both the accuracy of the point estimate and associated uncertainty. For many species, estimates of total annual catches did not improve by accounting for clustering of data, and in most cases led to increases in estimated uncertainty. However, we have demonstrated how a misleadingly optimistic view of precision leads to an increased risk of incorrectly concluding that unreported catches are significant (Figure 5).

The use of proxies such as encounter rate and Pearson's correlation coefficient between total catches and number of hooks to are useful tools for evaluating the performance of estimators across many species in a design-based framework. The Pearson's correlation coefficient is a fundamental factor of estimator performance for ratio estimators (Lohr, 2010) and we conclude based on best practices (ICES, 2007) that the application must be supported by evidence of a relationship rather than depending on assumptions. Encounter rates are considered for specific species in the USA's national bycatch reporting system (NMFS, 2011; Wigley et al., 2021) for which encounters are known to be rare. A similar approach is applied in Norwegian fisheries where seabirds (Bærum et al., 2019; Fangel et al., 2015) and marine mammals (Moan et al., 2020) are estimated independent to fish species (Berg and Nedreaas, 2020). These species groups are given more attention because of typically high conservation importance. However, defining empirical 'rules of thumb' are useful when many management decisions are made for species from a single estimation study. In the Barents Sea longline fishery, we have identified that estimators perform poorly when the Pearson's correlation coefficient falls below 0.25, and when rarer species have an encounter rate below 10 %. These findings can be used to guide future research to target future improvements of estimators, as well as for comparison with other fisheries to understand if estimator performance can be generalised for the Norwegian Reference Fleet or whether estimator performance is highly fishery-specific. For these comparisons, the testing methodology presented in this study will be a useful tool that can be quickly applied.

Representativeness is difficult to quantify in the direct context of unreported catches. Nevertheless, by evaluating the assumptions behind the estimators applied in this study, we identified that the cluster unit estimator may not be consistently suitable across species due to it being expected to overestimate total catches. The cluster ratio estimator mitigates against this overestimation by calculating the average catch per hook, which accounts for the higher fishing activity of Norwegian Reference Fleet vessels compared to the wider fleet (Figure 4). Nevertheless, we must still consider if data collected by the Norwegian Reference Fleet are representative of the fishery in terms of fishing strategy and catch composition. Although the ratio estimator accounts for the sampling of more active vessels, the sampled catch per unit effort (effort = hooks) may still not be representative, and different fishing strategies may mean sampled catch compositions are not representative. Clegg et al. (2022) found the Norwegian Reference Fleet tended to be representative of the wider fishery in relation to the reported component of total catches, particularly for commercial species, but identified a tendency to overestimate reported catches using Norwegian Reference Fleet data. This tendency is likely to also be applicable to the unreported component of catches that are estimated in this study. It is important to highlight that representativeness is discussed here in the specific context of estimating fishery-wide catches, and therefore cannot be directly applied to evaluating the representativeness of data with regards to temporal trends in catch per unit effort, or estimates of fish population parameters.

Unless there is specific knowledge on the sources of unreported catches, care must be taken when interpreting estimates to ensure that the correct course of action is taken. Unreported catches often suggest illegality caused by either discarding or intentional misreporting of landed catches, but there are many sources of unreported catches that are legal under discard policies. Discard bans typically come with exemptions such as non-quota species or high survivability (Borges et al., 2016; Catchpole et al., 2017; Karp et al., 2019). Low detail of catch reporting also leads to unreported catches. For example, difficulties in species identification of skate and rays (order: Rajiformes) leads catches being landed unidentified. Norwegian vessels are increasingly converting unwanted catches into fishmeal to increase utilisation of catches. However, vessels are not obliged to report the fishmeal ingredients with respect to relative contributions of individual species.

4.1. Improvements

The observations of total catches used in this study came solely from normal fishing activity by vessels in the Norwegian Reference Fleet. However, the Norwegian Directorate of Fisheries Monitoring and Surveillance Service (MSS) regularly hire fishing vessels to monitor catch rates in the Barents Sea fisheries, which may provide a supplementary data source to enhance estimates of unreported catches. For example, these data were used to map the bycatch risk and historical bycatch rates of cod in the Barents Sea shrimp fishery (Aldrin et al., 2011; Breivik et al., 2017, 2016). Whilst MSS data have been demonstrated to be suitable for this case study, we recommend a devoted study to address representativeness of these data before generalising the application to all fisheries for with the MSS cover.

This study has focused on determining the best estimator based on the current estimation framework in Norwegian Fisheries. The current practise for spatial stratification for estimating bycatches and discards in Norwegian fisheries is using statistical areas defined by the Norwegian Directorate of Fisheries. However, this stratification has not yet been optimised. The statistical area system is gridded, which is likely to explain some variations in catches such as latitudinal variations in the Barents Sea, but more complex drivers of spatial variations in catch composition such as temperature, depth or habitat are poorly described by a gridded statistical area system.

We estimated total catches of all species observed in the Barents Sea longline fishery, regardless of how often it was observed by the Norwegian Reference Fleet. Our study found that estimators tended to perform very poorly for the rarest of species. Future research can improve estimates of rare species in the design-based estimator framework. Using the delta lognormal estimator (Pennington, 1996) for example may help to improve estimator performance where assumptions are met. Performance could also be improved by applying model-based tools. On a single species basis, zero-inflated modelling can also improve both the bias and precision of parameter estimates (Martin et al., 2005). In a multispecies context, using a wider pool of information on the catch composition may help to explain the variations in catches of rare species (Thorson et al., 2016, 2015). This approach effectively ‘borrows’ information from more common species to predict the occurrence of rarer species.

In addition to the estimator bias and vessel sampling bias addressed in this study, we must also acknowledge biases in catch recording which could not be addressed in this study. The reliability of self-sampled data is subject to increased criticism (Kraan et al., 2013), given that the data will directly influence management decisions. The standard approach to quantifying reliability is through comparison with a data source of ‘known’ reliability (Roman et al., 2011). Such a study is yet to be done for the Norwegian Reference Fleet, but we argue based on values that the Norwegian Reference Fleet provide an overall reliable report of total catches. The programme is a trust-based collaboration, which is reflected in personal conversations and official meetings where fishers express their openness to cooperating with scientific research. Furthermore, fishers are paid under a contract to deliver high-quality data. Reliability is nevertheless dependent on the conservation status and management regulations, which differs between vessel groups, fishing gears, and species.

4.2. Generalisation

The Norwegian Reference Fleet includes vessels using a wide range of fishing gears in both coastal and offshore waters (Clegg and Williams, 2020). The specific sampling design for each fishing gear is adapted to account for unique gear characteristics. However, the ratio estimator is extendable to other gear-specific measures of fishing effort, given a strong correlation with total catches across vessels. The Norwegian Reference Fleet also has a coastal component, which again differs slightly to the sampling design for offshore vessels. Coastal fishing vessels in the Norwegian Reference Fleet do census reporting of total catches given the smaller scale of fishing activity. The cluster ratio estimator is nevertheless applicable, expect for the lack of need to estimate total catches per vessel (Equation 3.1), given that it is already known.

Coastal vessels in the Norwegian Reference Fleet explicitly record the discarded and retained portions of the total catch, allowing for a direct estimate of discarding. In 2019, the offshore Reference Fleet began a transition to recording retained catches, discards, and fishmeal explicitly, rather than a single value for total catch. Direct observations of discards remove the need to infer unreported catches through a comparison with reported catches. The cluster ratio estimator will still be applicable in the context of discards, given that there is a strong relationship with the chosen measure of fishing effort.

4.3. Conclusions

We conclude that cluster-based estimators should be used when estimating unreported catches based on data collected by the Norwegian Reference Fleet. The difficulties in evaluating the representativeness of data, coupled with variable relationship between total catches and fishing effort (number of hooks), means that it is not possible to conclude a single best estimator for all species. However, there are clear indications that the cluster unit estimator is not unbiased as theoretically expected due to nonrepresentative sampling of vessels that are more active in the fishery. Whilst there are biases identified with the cluster ratio estimator, they are identifiable by evaluating the relationship between total catches and fishing effort. We therefore recommend based on best practice methodology (ICES, 2007) that the unit and ratio estimator are applied and compared. If large differences are found, then further investigations can identify the reasons. Annual total catch estimates of rare species with an encounter rate below 10 % or Pearson’s correlation coefficient below 0.25 should be interpreted with caution, and we suggest further development of methods to reduce biases and improve precision. In this future research, the testing methodology presented in this study will be a useful tool for comparing across estimators and fisheries.

Acknowledgements
We would like to thank the fishers participating in the Norwegian Reference Fleet, past and present, for contributing data and knowledge to this study. We also thank Kjell Nedreaas and Geir Blom for their comments on an earlier draft of the manuscript.

Funding
This work was funded equally by the Norwegian Institute of Marine Research and the Norwegian Directorate of Fisheries.

Conflict of interest
The authors declare no conflict of interest.

Data availability statement
The data underlying this article cannot be shared publicly due the sensitivity of the contents and the privacy of fishers involved in data collection. The data will be shared on reasonable request to the corresponding author.

References
Aanes, S., Pennington, M., 2003. On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. ICES J. Mar. Sci. 60, 297–303. https://doi.org/10.1016/S1054-3139(03)00008-0

Aldrin, M., Mortensen, B., Storvik, G., Nedreaas, K., Aglen, A., Aanes, S., 2011. Improving management decisions by predicting fish bycatch in the Barents Sea shrimp fishery. ICES J. Mar. Sci. 69, 64–74. https://doi.org/10.1093/icesjms/fsr172

Bærum, K.M., Anker-Nilssen, T., Christensen-Dalsgaard, S., Fangel, K., Williams, T., Vølstad, J.H., 2019. Spatial and temporal variations in seabird bycatch: Incidental bycatch in the Norwegian coastal gillnet-fishery. PLoS One 14, e0212786. https://doi.org/10.1371/journal.pone.0212786

Batsleer, J., Hamon, K.G., van Overzee, H.M.J., Rijnsdorp, A.D., Poos, J.J., 2015. High-grading and over-quota discarding in mixed fisheries. Rev. Fish Biol. Fish. 25, 715–736. https://doi.org/10.1007/s11160-015-9403-0

Berg, H.S.F., Nedreaas, K., 2020. Estimering av utkast i Norsk kystfiske med garn 2012-2018. Institute of Marine Research, Brage IMR.

Borges, L., Cocas, L., Nielsen, K.N., 2016. Discard ban and balanced harvest: a contradiction? ICES J. Mar. Sci. 73, 1632–1639. https://doi.org/10.1093/icesjms/fsw065

Borges, L., Zuur, A.F., Rogan, E., Officer, R., 2005. Choosing the best sampling unit and auxiliary variable for discards estimations. Fish. Res. 75, 29–39. https://doi.org/10.1016/j.fishres.2005.05.002

Breivik, O.N., Storvik, G., Nedreaas, K., 2017. Latent Gaussian models to predict historical bycatch in commercial fishery. Fish. Res. 185, 62–72. https://doi.org/10.1016/j.fishres.2016.09.033

Breivik, O.N., Storvik, G., Nedreaas, K., 2016. Latent Gaussian models to decide on spatial closures for bycatch management in the Barents Sea shrimp fishery. Can. J. Fish. Aquat. Sci. 73, 1271–1280. https://doi.org/10.1139/cjfas-2015-0322

Catchpole, T.L., Ribeiro-Santos, A., Mangi, S.C., Hedley, C., Gray, T.S., 2017. The challenges of the landing obligation in EU fisheries. Mar. Policy 82, 76–86. https://doi.org/10.1016/j.marpol.2017.05.001

Clegg, T., Williams, T., 2020. Monitoring bycatches in Norwegian fisheries - Species registered by the Norwegian Reference Fleet (No. Rapport fra havforskningen; 2020-8).

Clegg, T.L., Fuglebakk, E., Ono, K., Vølstad, J.H., Nedreaas, K., 2022. A simulation approach to assessing bias in a fisheries self-sampling programme. ICES J. Mar. Sci. 79, 76–87. https://doi.org/10.1093/icesjms/fsab242
Diamond, B., Beukers-Stewart, B.D., 2011. Fisheries discards in the North Sea: waste of resources or a necessary evil? Rev. Fish. Sci. 19, 231–245. https://doi.org/10.1080/10641262.2011.585432
Dingsør, G.E., 2001. Estimation of discards in the commercial trawl fishery for cod. Dep. Fish. Mar. Biol. Masters thesis, University of Bergen.
Efron, B., Tibshirani, R.J., 1994. An Introduction to the Bootstrap. CRC Press.
Fangel, K., Aas, Ø., Vølstad, J.H., Bærum, K.M., Christensen-Dalsgaard, S., Nedreaas, K., Overvik, M., Wold, L.C., Anker-Nilssen, T., 2015. Assessing incidental bycatch of seabirds in Norwegian coastal commercial fisheries: Empirical and methodological lessons. Glob. Ecol. Conserv. 4, 127–136. https://doi.org/10.1016/j.gecco.2015.06.001
Fernandes, A.C., Oroszlányová, M., Silva, C., Azevedo, M., Coelho, R., 2021. Investigating the representativeness of onboard sampling trips and estimation of discards based on clustering. Fish. Res. 234. https://doi.org/10.1016/j.fishres.2020.105778
Gezelius, S.S., 2006. Monitoring fishing mortality: Compliance in Norwegian offshore fisheries. Mar. Policy 30, 462–469. https://doi.org/10.1016/j.marpol.2005.06.004
Gilman, E., Perez Roda, A., Huntington, T., Kennelly, S.J., Suuronen, P., Chaloupka, M., Medley, P.A.H., 2020. Benchmarking global fisheries discards. Sci. Rep. 10, 1–8. https://doi.org/10.1038/s41598-020-71021-x
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. https://doi.org/10.1093/icesjms/fst094
Gullestad, P., Blom, G., Bakke, G., Bogstad, B., 2015. The “Discard Ban Package”: Experiences in efforts to improve the exploitation patterns in Norwegian fisheries. Mar. Policy 54, 1–9. https://doi.org/10.1016/j.marpol.2014.09.025
Hønneland, G., 2000. Compliance in the Barents Sea Fisheries. How fishermen account for conformity with rules. Mar. Policy 24, 11–19. https://doi.org/10.1016/S0308-597X(98)00058-X
ICES, 2021. Arctic Fisheries Working Group 2021 report. https://doi.org/10.17895/ICES.PUB.8196
ICES, 2007. Report on the workshop on discard raising procedures. ICES WKDRP CM 2007/ACFM:06 Ref RMC PGCCDBS. San Sebastian, Spain.
ICES, 2003. Workshop on discard sampling methodology and raising procedures. ICES Planning Group on Commercial Catch, Discards and Biological Sampling, Chalottenlund, Denmark 2-4 September 2003.
Karp, W.A., Breen, M., Borges, L., Fitzpatrick, M., Kennelly, S.J., Kolding, J., Nolde Nielsen, K., Viðarsson, J.R., Cocas, L., Leadbitter, D., 2019. Strategies Used Throughout the World to Manage Fisheries Discards – Lessons for Implementation of the EU Landing Obligation, in: Uhlmann, S.S., Ulrich, C., Kennelly, S.J. (Eds.), The European Landing Obligation: Reducing Discards in Complex, Multi-Species and Multi-Jurisdictional Fisheries. Springer, Cham, pp. 3–26. https://doi.org/10.1007/978-3-030-03308-8
Kelleher, K., 2005. Discards in the world’s fisheries. An update (No. FAO Fisheries Technical Paper. No. 470.). Rome, FAO.
Kennelly, S.J., 2020. Bycatch Beknown: Methodology for jurisdictional reporting of fisheries discards – Using Australia as a case study. Fish Fish. 21, 1046–1066. https://doi.org/10.1111/faf.12494
Kraan, M., Uhlmann, S., Steenbergen, J., van Helmond, A.T.M., van Hoof, L., 2013. The optimal process of self-sampling in fisheries: lessons learned in the Netherlands. J. Fish Biol. 83, 963–973. https://doi.org/10.1111/jfb.12192

Lohr, S.L., 2010. Sampling: Design and Analysis, 2nd ed. Chapman and Hall/CRC.
Martin, T.G., Wintle, B.A., Rhodes, J.R., Kuhnert, P.M., Field, S.A., Low-Choy, S.J., Tyre, A.J., Possingham, H.P., 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecol. Lett. 8, 1235–1246. https://doi.org/10.1111/j.1461-0248.2005.00826.x
McBride, M.M., Fotland, Å., 1996. Estimation of unreported catch in a commercial trawl fishery. J. Northwest Atl. Fish. Soc. 18, 31–41.
Moan, A., 2016. Bycatch of harbour porpoise, harbour seal and grey seal in Norwegian gillnet fisheries. Masters thesis, University of Oslo.
Moan, A., Skern-mauritzen, M., Vølstad, J.H., Bjørge, A., 2020. Assessing the impact of fisheries-related mortality of harbour porpoise (*Phocoena phocoena*) caused by incidental bycatch in the dynamic Norwegian gillnet fisheries. ICES J. Mar. Sci. 77, 3039–3049. https://doi.org/10.1093/icesjms/fsaa186
Nedreaas, K., Iversen, S., Grethe, K., 2015. Preliminary estimates of total removals by the Norwegian marine fisheries, 1950-2010. Working Paper #2015-94, Fisheries Centre, University of British Columbia, Vancouver, Canada.
Nelson, G.A., 2014. Cluster Sampling: A Pervasive, Yet Little Recognized Survey Design in Fisheries Research. Trans. Am. Fish. Soc. 143, 926–938. https://doi.org/10.1080/00028487.2014.901252
NMFS, 2011. U.S. National Bycatch Report, U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-117E, 508 p. U.S. Department of Commerce, NOAA Technical Memo NMFS-F/SPO-117E, 508 p.
Norwegian Directorate of Fisheries, 2021. Norske Omregningsfaktorer - For omregning av landet produktvekt av marin fisk og andre marine arter til rund vekt - for fiske i det nordlige Atlanterhavet og i andre farvann.
Pennington, M., 1996. Estimating the mean and variance from highly skewed marine data. Biometrics 39, 281–286. https://doi.org/10.2307/2530830
Pérez Roda, M.A., Gilman, E., Huntington, T., Kennelly, S.J., Suuronen, P., Chaloupka, M., Medley, P., 2019. A third assessment of global marine fisheries discards. FAO Fisheries and Aquaculture Technical Paper No. 633. Rome, FAO. 78 pp.
Pitcher, T.J., Watson, R., Forrest, R., Þo¤rValty¤sson, H., Gue¤nette, S., 2002. Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish Fish. 3, 317–339. https://doi.org/10.1046/j.1467-2979.2002.00093.x
Roman, S., Jacobson, N., Cadrin, S.X., 2011. Assessing the reliability of fisher self-sampling programs. North Am. J. Fish. Manag. 31, 165–175. https://doi.org/10.1080/02755947.2011.562798
Thorson, J.T., Ianelli, J.N., Larsen, E.A., Ries, L., Scheuerell, M.D., Szuwalski, C., Zipkin, E.F., 2016. Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. Glob. Ecol. Biogeogr. 25, 1144–1158. https://doi.org/10.1111/geb.12464
Thorson, J.T., Scheuerell, M.D., Shelton, A.O., See, K.E., Skaug, H.J., Kristensen, K., Warton, D., 2015. Spatial factor analysis: a new tool for estimating joint species distributions and correlations in species range. Methods Ecol. Evol. 6, 627–637. https://doi.org/10.1111/2041-210x.12359
Wigley, S.E., Asci, S., Benjamin, S., Chamberlain, G., Cierpich, S., Didden, J., Drew, K., Legault, C., Linden, D., Murray, K., Potts, D., Sampson, K., Tholke, C., 2021. Standardized bycatch reporting methodology 3-year review report - 2020: Reviewing SBRM years 2018, 2019, and 2020 (No. NOAA Technical Memorandum NMFS-NE-266).

PAPER IV

Title

Discards of cod (*Gadus morhua*) in the Norwegian coastal fisheries: improving past and future estimates

Authors

Hilde Sofie Fantoft Berg*[1,2], Thomas Laurence Clegg*[1], Geir Blom[2], Jeppe Kolding[3], Kotaro Ono[1], Kjell Nedreaas[1]

* These authors contributed equally to the study.

1. Institute of Marine Research/Havforskningsinstituttet, Postboks 1870 Nordnes, NO-5817 Bergen, Norway

2. Directorate of Fisheries, PO Box 185, 5804 Bergen, Norway

3. University of Bergen, P.O. Box 7800, NO-5020, Bergen, Norway

Corresponding author: tom.clegg@hi.no +47 55 23 85 00 (TL Clegg)

Abstract

Discarding can be an unknown source of biases and uncertainties in stock assessments. Discarding patterns and quantities vary so a routine methodology for estimating discards is important to give a better picture of total catches, and potentially mortality, in fisheries. Using data from the Norwegian Reference Fleet between 2012 and 2018, this study presents a revised methodology for estimating discards of cod (*Gadus morhua*) in the Norwegian coastal gillnet fisheries which accounts for variations in discarding between vessels and uncertainties in the conversion of numbers to weight discarded. The estimated average discard rate of cod (weight of cod discarded as percentage of total weight caught) is 0.55 % (95 % CI: 0.45-0.70 %), although discard rates in southern areas were an order of magnitude higher than in northern areas. We also present an exploratory analysis of the drivers behind discarding using a random forest regression model. Spatial variations and fishing intensity were identified as the most important drivers of discarding. Results from this study suggest ways in which self-sampled data can be used to estimate discards in Norwegian coastal fisheries, and where accuracy of future estimates can be improved when a higher resolution data collection programme is established.

Introduction

Discarding of fish at sea is widely perceived as an unethical waste of resources and discarding can be a major unknown factor in stock assessments. Discarding patterns can vary greatly between different fisheries, areas, gears, and target species, but an estimation of these variations is sometimes complicated by limited data availability. Simulation studies have explored the consequences of ignoring discards in stock assessments (Dickey-Collas *et al.*, 2007; Perretti *et al.*, 2020), as well as accounting for trends in discarding over time to ensure accurate estimates of the fishery status (Rudd and Branch, 2017; Cook, 2019).

Incentives to discarding are often based on conflicting regulatory and economic factors. Examples include catch of unmarketable or undersized fish, ‘choking’ when a vessel has exceeded the quota of targeted marketable catch, ‘high-grading’ when fish of lower value is discarded so that fishing for more valuable catch can continue (Rochet and Trenkel, 2005; Batsleer *et al.,* 2015; Karp *et al.*, 2019), or catch of fish that is unfit for human consumption (e.g., disease, gear damage, or scavenging). Discarding is illegal in Norway, but with some exemptions (e.g., viable fish can be released back to the sea) (Gullestad *et al.*, 2015). The scale of discarding has largely been unknown in Norwegian fisheries since the discard ban was implemented in 1987 (Gullestad *et al*. 2015; Karp *et al.*, 2019). As a result, Norway does not currently provide discard information for cod (*Gadus morhua*) stock assessments in either northeast Arctic, Norwegian coastal (ICES, 2020a) or North Sea (ICES, 2021) stocks and therefore discards are currently assumed to be negligible.

Independent scientific observers are widely seen as the most reliable data collection method for discarding (Pérez-Roda *et al*. 2019). However, this approach becomes unreliable under a discard ban where the presence of an observer may deter fishers from discarding (Benoît and Allard, 2009), especially if the observer must report illegal activities or if there is less than 100 % observer coverage (Ewell *et al.*, 2020). To address this issue, countries are increasingly moving towards self-reported data to support or replace observer programmes (Mangi *et al*., 2013). The Norwegian Institute of Marine Research (IMR) monitors discarding using the Norwegian Reference Fleet, an enhanced self-sampling programme in which participating vessels are paid to provide detailed information on catches and fishing activity regularly and confidentially. Data from the Norwegian Reference Fleet have previously been used to estimate bycatches of fish (Berg and Nedreaas, 2020), seabirds (Fangel *et al*. 2015; Bærum *et al.* 2019) and harbour porpoises (Moan *et al*., 2020) in the coastal gillnet fisheries.

In coastal fisheries, the Norwegian Reference Fleet record a wide range of variables which cannot be incorporated into current estimators due to the limited comparable data submitted by the rest of the vessels in the fishery. However, the Norwegian authorities have approved an extension of requirements for daily electronic reporting system for coastal vessels, which will be gradually implemented between 2022 and 2024, providing higher resolution data on catches and fishing activity. This creates the opportunity to explore the complex drivers of discarding in more detail, based on information recorded by the Norwegian Reference Fleet, and evaluate whether the electronic reporting system dataset can be useful for improving estimations of discarding in the future.

In this study, we present a development in the methodology for estimating discards in the Norwegian coastal gillnet fisheries between 2012 and 2018, which accounts for variations in discarding behaviour between vessels. Of particular importance to these estimations is the clustering of samples in the Norwegian Reference Fleet sampling design. The importance of clustering has been discussed in many previous studies (Aanes and Pennington, 2003; Helle and Pennington, 2004; Pennington and Helle, 2011; Clegg *et al*., 2021), and is currently accounted for when using generalised linear models to estimate discards (Bærum *et al.* 2019; Moan *et al*., 2020). However, traditional applications of a design-based estimator in Norwegian fisheries do not account for variability in discarding between vessels, which may be resulting in an overestimation of precision (Lohr, 2010; Nelson, 2014). Importantly, the scale of this impact is unknown until a clustered estimator is applied. The study also presents an exploratory analysis of the drivers behind the discarding of the Norwegian Reference Fleet using a random forest regression model, and in preparation for improvements in the mandatory catch reporting system. This aims to identify important potential drivers behind discarding behaviour in the coastal fisheries and suggest a framework for a model-based approach to estimating discards once the more detailed electronic reporting system is fully operational.

Case study fishery

Our study focuses on Norwegian coastal gillnet fisheries, which we define as vessels under 15 m LOA using gillnets within 12 nautical miles of the coast (Figure 1). There is an important division in ecosystems and stock distributions at 62 °N, and this line subsequently represents a division in Norwegian fisheries management and regulations. Between 2012 and 2018, commercial vessels under 15 m LOA accounted for 33 % of total reported cod catches by Norwegian vessels. Within the coastal fisheries, gillnets accounted for 59 % of reported cod catches by commercial vessels under 15 m LOA in coastal statistical areas. Hook (longline and jigging) and Danish seine fisheries accounted for 34 % and 8 % of cod catches, respectively. However, these latter fisheries were excluded from the study as the Norwegian Reference Fleet programme prioritises these fishing gears only for specific areas and target species (Clegg and Williams, 2020).

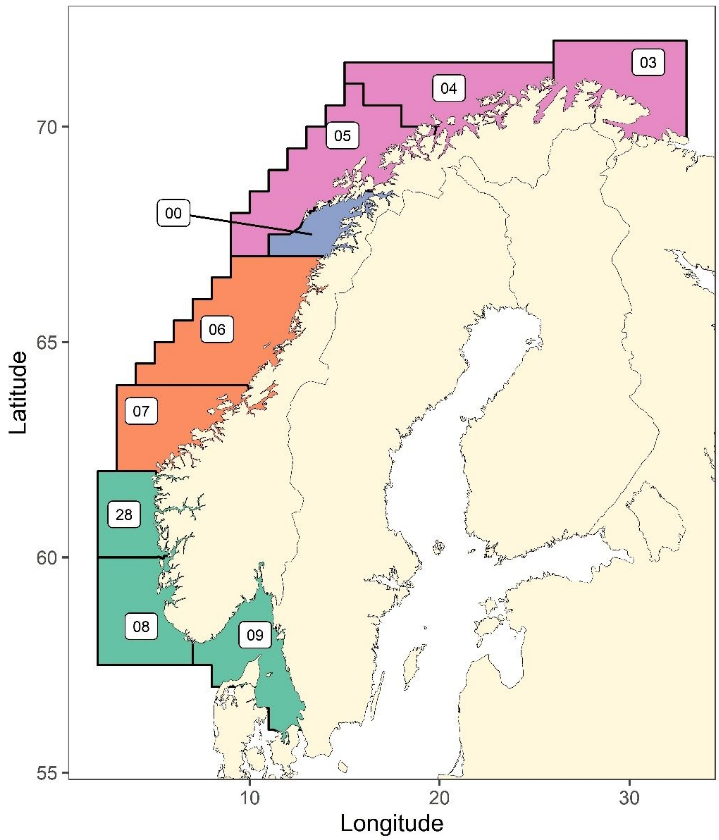


Figure 1. Map of study area including statistical areas defined by the Norwegian Directorate of Fisheries. Shading indicates regions used in estimation procedure. The division between north and south management systems is at 62 °N (the boundary between statistical areas 07 and 28). There was a small change in the geographic extent of areas 08 and 09 from 2018. This is accounted for in the analyses, but not shown in the maps presented.

Catch of cod in the Norwegian coastal areas is a combination of Northeast Arctic cod, Norwegian coastal cod, and North Sea cod. Between January and April, Northeast Arctic cod migrate for annual spawning from the Barents Sea to the Norwegian coast, mainly areas 00, 05 and 06. The Norwegian coastal cod spends its whole life along the Norwegian coast, in fjords and coastal sea banks. In southern parts of Norway North Sea cod occasionally migrate to coastal areas. The three stocks overlap to varying degree between both seasons and areas (ICES, 2020a).

In the coastal areas of Norway, cod is for the most part caught in three gillnet fisheries. The targeted cod fishery, the mixed gadoid fishery, and the anglerfish (*Lophius piscatorius*) fishery. The targeted cod fishery targets the Northeast Arctic cod on their spawning migration and concentration in the Lofoten area in the first quarter of the year. The mixed gadoid fishery operates throughout the year, and in addition to cod target species like saithe (*Pollachius virens*), haddock (*Melanogrammus aeglefinus*), pollack (*Pollachius pollachius*), and ling (*Molva molva*). The targeted cod and mixed gadoid fisheries use similar mesh sizes, from a minimum of 156 mm to approximately 210 mm stretched mesh. It is therefore difficult to differentiate between these two fisheries besides assumptions based on area and season, even if mesh sizes are known. The targeted anglerfish fishery which operates throughout the year except between 1st March to 20th May between 62-64 ˚N, and except 20th December to 20th May north of 64 ˚N, uses gillnets with mesh sizes > 360 mm stretched mesh. The anglerfish fishery overlaps with the mixed gadoid fishery in both space and time. The catches of cod are generally lower in the south than in the north for all gillnet fisheries. The three cod stocks have separate assessments and management-plans, and the quotas are determined in annual negotiations between the relevant countries. There is insufficient information available in data collected by the Norwegian Reference Fleet to determine the stock origin of individual discards and we therefore do not differentiate between individual stocks in this study.

Under the Marine Resources Act 2008, discarding of all species is in principle illegal, and all catches must be landed and reported. However, the “Discard Ban Package” (see Gullestad *et al.*, 2015 for detailed description) contains several exemptions and measures to ease the discard ban with aims of making it more practical to follow. Firstly, this includes formal exemptions from the ban for fish which are alive when released, and informal exemptions for damaged fish unfit for human consumption. Secondly, other measures include compensation for landing of some unwanted catches which could otherwise end up as discards; an obligation to move away from areas with high levels of illegal catches, such as undersized fish; requirements for selectivity devices in certain fishing gear; and adjustments to the quota system to include a certain amount of bycatch to reduce discarding incentives. Despite these additional measures to avoid unwanted catches and incentivise their landing if incurred, the risk of illegal discarding should still be acknowledged. The Norwegian Coast Guard enforce fishing regulations, including the discard ban, through at-sea surveillance, and inspections (on all vessels from all nations), whilst the Fisheries Directorate run both at-sea surveillance (only on Norwegian vessels) and shore-based inspections of landings and sales.

Data

The Sales Note database

The reporting system in the coastal gillnet fishery is centred around the landing and sale of catches. All Norwegian catches are sold through registered sales organisations, for which there were six in the study period (reduced to five since 2020). The sales organisations are responsible for correct landing statistics, deducting quota, compensating the landing of unintended catches, and reporting any suspected illegal activity. Skippers are required by law to report first-hand sale of catches (Norwegian Ministry of Trade, Industry and Fisheries, 2014), which are signed by both the seller and buyer. This sales note database therefore creates a census of all landed catches by species and weight in Norwegian waters. When a vessel returns to port to land catches, they must submit a landing note which includes the total catch weight of each species, statistical area of catch, and date it was landed. For each sale of fish, a sales note is generated which reports the quantity sold.This quantity is then deducted from the associated landing note and from the vessels quota. Sales notes are not a reliable metric of fishing effort because multiple sales notes can be generated for one catch depending on the number of buyers. Sales notes therefore need to be back traced and aggregated to individual trips based on the landing date reported on the sales notes. Coastal vessels operate on day trips, meaning that one reported landing date should generally represent one day of fishing. However, we expect some variability in this assumption due to complex sales of fish from multiple trips, delayed reporting, or due to reporting errors. To evaluate whether landing date is a suitable identifier of trips, we linked the daily observations from vessels in the Norwegian Reference Fleet to the most recent landing date following each observation. This linkage determined that 75 % of trips comprise of one fishing day, and 98 % of trips comprised of three fishing days or less. Comparing this to the larger variabilities in trip duration and associated catches in offshore fisheries, we concluded that landing dates are a suitable identifier of fishing trips.

The Norwegian Reference Fleet data

The participating vessels in the Norwegian Reference Fleet are selected through an open tender process and are paid for high resolution self-sampling and recording of catches. The tender specifications (see Clegg and Williams, 2020) aim to select vessels that are representative of the wider fleet in each statistical area. If multiple vessels meet the required specifications, then the contract is awarded randomly. Each vessel has a contact person employed at IMR that follow up and regularly visits the vessels to guide methods and procedures for correct sampling protocols. The accuracy and reliability of self-sampled discard data is a recognised concern (Kraan *et al.*, 2013), given that data could be used for prosecution, and results could affect fishery access. There is an agreement between fishers, scientists, and the Norwegian authorities that data shall not be used for prosecution. To date, this agreement has not been compromised, which provides fishers with the trust to record discards with the assurance that the data shall only be used for scientific purposes. Lastly, misreporting is mitigated by a willingness to participate and honest communication between fisheries and scientists. Furthermore, a lot of effort and emphasis is invested in the Norwegian Reference Fleet programme to ensure true and correct sampling and reporting.

This study uses data from the Coastal Reference Fleet, a subdivision of the Norwegian Reference Fleet for vessels under 15 m LOA. Coastal vessels record catches and fishing activity for every calendar day they are active, which we refer to in this study as a fishing operation. This includes retained catches (recorded as weight) and discards (recorded as numbers) by species. Fishers do not record whether discards are legal (e.g., viable or damaged) or illegal. To target a wider range of species, skippers often have gillnets with different specifications (e.g., mesh size, material) set in different locations. To account for this behaviour, the sampling guidelines specify that if groups of gillnets differ significantly in specifications and geographic locations, then these should be recorded as separate fishing operations. In addition to daily reporting of catches and discards, a representative sample of 20 fish per species are taken each week from each of the retained and discarded portions of the catches for length measurements.

Statistical analyses

Data handling and statistical analyses were done in R (version 4.1.0; R Core Team, 2021).

Defining the study fisheries in datasets

The sales note database only classifies fishing gears by broad groups. This makes it possible to distinguish between fisheries using different gear types, such as gillnets, hooks, or trawls, but there is no detailed information on the specifications of these gears, which for gillnet fisheries would be mesh size, materials, soaking time, and number of nets used. Therefore, excluding non-gillnet fisheries from this study must be based on the limited information available.

The pelagic gillnet fisheries have a low bycatch rate for cod relative to the demersal gillnet fisheries where cod is more likely to encounter the fishing gear. We therefore excluded pelagic fisheries from the study by removing trips in which a pelagic species contributed the largest proportion to reported catch weight.

Recreational fishing with gillnets is popular along the Norwegian coast. Whilst there is no obligation to report recreational catches, any catches that are sold (limited to 50 000 NOK per year) must be reported using a sales note. As this study is limited to commercial fisheries, we excluded recreational catches from the study by removing sales notes without a documented vessel length, which was deemed as the best identifier of recreational vessels.

Estimating total discards

To estimate total discards in the coastal gillnet fisheries, we first needed to standardise the sample (Norwegian Reference Fleet) and the population (sales note database) datasets. We therefore aggregated all fishing operations by Norwegian Reference Fleet vessels within the study to summarise total discards per trip. Fourteen fishing days could not be associated with a landing date, due to data recording errors, so we assumed these were one day trips. For the same reason, we also assumed that trips comprising more than five fishing days were erroneous, resulting in the removal of 31 trips consisting of 273 fishing days. After this initial cleaning step, we had a dataset containing 6 662 trips from 43 Norwegian Reference Fleet vessels over the study period (Table 1).

Table 1. Number of sampled vessels and trips with detailed information on fishing activity from the Norwegian Reference Fleet (NRF) compared to the whole Norwegian fleet within coastal gillnet fisheries (vessels < 15 m LOA) in the period 2012-2018. Number of vessels across all years (last row) is not the sum of individual years because vessels are active across multiple years.

Year	Number of vessels			Number of trips		
	NRF	Whole fleet	% Sampled	NRF	Whole fleet	% Sampled
2012	20	2 273	0.9	729	62 917	1.2
2013	16	2 054	0.8	939	57 458	1.6
2014	16	1 942	0.8	908	58 217	1.6
2015	20	1 943	1.0	857	51 901	1.7
2016	22	1 965	1.1	1 156	53 465	2.2
2017	20	1 992	1.0	1 028	52 601	2.0
2018	20	2 081	1.0	1 045	56 903	1.8
All years	43	3 497	1.2	6 662	393 462	1.7

The estimation methodology is based on Berg and Nedreaas (2020), who used a stratified unit estimator (Lohr, 2010). However, the estimator was redefined to reflect the clustered sampling routine of the Norwegian Reference Fleet, which defines vessels as the primary sampling unit from which fishing operations are repeatedly sampled and provide a more accurate estimate of variance.

This was defined as a ratio estimator based on the assumption that total discards are positively correlated with the number of trips for individual vessels (Lohr, 2010; r(41) = 0.78, p < 0.001). For each sample stratum (defined as a combination of statistical area (Figure 1), annual quarter, and year; n = 258), the total number of discarded cod $\hat{Y}$was estimated by:

$$\hat{Y} = M_0 \frac{\sum_{i=1}^{n} \sum_{j=1}^{m_i} M_i \frac{y_{ij}}{m_i}}{\sum_{i=1}^{n} M_i} \quad (1)$$

where for trip j by vessel i, y is the number of cod discarded; m is the number of trips sampled and n is the number of vessels sampled; M_i is the total number of trips by sampled vessel i and M_0 is the total number of trips in the population. This estimator assumes that sampled vessels are a simple random selection from the wider fishing fleet. Note that because coastal vessels record discards for all trips (i.e., $m_i = M_i$), Equation 1 simplifies down to the stratified unit estimator used by Berg and Nedreaas (2020).

Strata with fewer than two sampled trips were defined as unsampled (n = 51). To estimate discards in unsampled strata, observations were borrowed from adjacent statistical areas in the same period, by assuming that discarding behaviour is more similar across statistical areas in the study than across quarters and years. For each unsampled stratum, the areas included were incrementally expanded for imputation, whilst keeping fixed the annual quarter and year, until there were sufficient samples for estimating discards. This meant imputing based on the mean across all areas in (1) the region (Figure 1), then (2) management system (north or south of 62 °N latitude), and finally (3) all areas in the study.

Estimating the discard rate

The discard rate of cod, as the percentage of total catch of cod in weight, is defined as:

$$Discard\ rate(\%) = 100 \times (\frac{discards}{landings + discards}) \quad (2)$$

Because information on total landed cod is only available in weight, the total weight of discarded cod needed to be estimated. This is only possible by conversion, as the Norwegian Reference Fleet do not record weights of individual fish. For each year, annual quarter, and management system (north/south of 62 °N), we produced a length-weight relationship (Equation 3) by using data from all fisheries-dependent and -independent sampling programmes in the study fishery for which IMR has access (summary of data sources available in Appendix A). Parameters *a* and *b* in the length-weight relationship were estimated using nonlinear least squares.

$$W = aL^b \quad (3)$$

Where W and L are weight (kg) and length (cm) respectively. We averaged these estimated weights per stratum then multiplied values by the estimated number (Equation 1) to produce an estimate of total weight of cod discarded and subsequently estimated discard rate using Equation 2.

Variance in discard estimates

We estimated the 95 % confidence interval (CI) of discard estimates using the bootstrapping method (Efron and Tibshirani, 1994). We present a refined procedure for estimating variance in total catches which reflects the clustered sampling design of the Norwegian Reference Fleet by accounting for the variation between vessels. For each bootstrap replicate, vessels were resampled in each year, then estimated discards using Equation 1 and the defined imputation procedure. As sampling of coastal vessels follows a one-stage cluster sampling design (i.e., all trips are sampled for each vessel), total discards per vessel were known, so trips were not resampled.

We accounted for uncertainty in the conversion of numbers discarded to biomass in both the selection of fishing operations for length measurements and the length-weight relationship. Firstly, we resampled with replacement the discarded fish sampled weekly for length measurements by each vessel. Secondly, to account for variation in the length-weight relationship we performed a parametric bootstrap of model parameters using the fitted model.

To evaluate the importance of accounting for additional variance in the conversion from number to weight, we compared the coefficient of variation of estimated discard rates for individual strata before and after including the additional sources of variation.

Modelling important drivers of discards

To identify important potential drivers of discarding in the coastal cod fishery, we fitted a random forest regression model (Breiman, 2001). We chose a random forest model over a generalised linear or additive modelling (GLM/GAM) framework due to the minimal assumptions needed to understand the complex reasons for discarding, allowing for a more explorative analysis. Random forests require no assumption of relationships between discarding and the explanatory variables, nor interactions between explanatory variables. Furthermore, the inclusion of multiple potentially uninfluential and collinear variables has little impact on model fit (Cutler, 2007).

Relationships between the response and explanatory variables in a GLM or GAM are typically evaluated using the strength and statistical significance of the relationship, which are strongly influenced by prior assumptions and decisions made in the model selection procedure (Burnham and Anderson, 2002). Contrastingly, random forests produce a useful measure of variable importance, which helps to identify the most important explanatory variables. We calculated variable importance using the permutation method (Breiman, 2001). After calculating the prediction accuracy of the fitted model, each variable is randomly permuted in turn and the predictions re-calculated. The importance is calculated as the mean decrease in model accuracy, scaled by the standard error. If an explanatory variable is strongly associated with discards of cod, then model accuracy will decrease when the values are permuted.

For the modelling of important drivers of discards, we used observations of total number of discarded cod for individual fishing operations by the Norwegian Reference Fleet, which allows us to include more detailed information on gear specifications and geographic location. The full list of explanatory variables used in the model are listed in Table 2. These variables were selected from the available data to capture the complex interaction of factors affecting discarding behaviour based on species distribution, fishing behaviour, and management regulations. Some studies have noted possible downsizing of importance for strongly correlated variables (Boulesteix *et al.*, 2012). In this study, only latitude and longitude had a correlation > 0.7 amongst the continuous variables (Figure S2; Supplementary Materials). Correlation is also expected between categorical descriptions of spatial and temporal variations in the model (e.g., month and quarter; statistical area and management system). We decided to keep these variables in the model as relative importance will still be interpretable and inform future predictive models. Where two spatial or temporal variables are important, then devoted methods are available (Yan *et al.*, 2021), and the decision of which variable to use will also be driven by data availability. Geographic coordinates were missing for 35 observations, which were imputed with the centroid of the reported statistical location (a sub-division of statistical area typically spanning one degree of longitude and half a degree of latitude). Fourteen observations with missing soak time values were removed from the study as soak times vary too much to assume an imputed value. Fishing depth was imputed for 137 observations with the recorded depth from the geographically nearest observation. All imputed values were from within 4 km, and 23 were from an observation with the same geographical coordinates. Weekly prices per statistical area were deemed erroneous if they were outside 1.5 times the interquartile range of all prices and were imputed with the most recent price prior to it in that statistical area. Of the 2 938 weekly price values, 29 were imputed from the previous week and 8 were imputed from between 2-4 weeks prior. This data cleaning procedure resulted in a dataset containing 10 090 fishing operations.

Table 2. Explanatory variables included in the random forest model to predict variations in discarding of cod in the Norwegian coastal gillnet fisheries. Random variables are included to detect bias in importance in the random forest model. Letters in parentheses refer to grouping of explanatory variables: a = species distribution, b = fishing behaviour, and c = management

Variable	**Type**	**Description**
Year (abc)	Factor	2012-2018
Month (abc)	Factor	Calendar month
Quarter (abc)	Factor	Calendar quarter
Vessel (b)	Factor	Unique vessel identifier (call signal)
Latitude (abc)	Continuous	Decimal degrees north
Longitude (abc)	Continuous	Decimal degrees east
Depth (ab)	Integer	Maximum fishing depth (nearest metre)
Number of nets (bc)	Integer	Total number of nets
Mesh size (bc)	Factor	Five categories: <140, 140-180, 181-260, >260 mm, mixed
Soak time (b)	Integer	Soak time of nets (hours)
Statistical area (c)	Factor	Management area defined by the Norwegian Directorate of Fisheries (Figure 1)
Management system (c)	Factor	Division between two fisheries management systems (north or south of 62 °N latitude)
Landed weight of cod (b)	Continuous	Total weight of retained cod (tonnes)
Target species (b)	Factor	Species contributing most to total retained catch weight: cod, anglerfish, other
Cod price (b)	Continuous	Average weekly prices of cod (NOK/kg) in each statistical area. Prices included sales of fresh fish sold either whole, gutted, or headed and gutted. All prices standardised to the 2018 consumer price index.
Random variable 1	Continuous	Random values from a normal distribution with mean = 0 and standard deviation = 1
Random variable 2	Continuous	Random values from a uniform distribution between 0 and 1

Random variable 3	Factor	Random factor with 5 levels
Random variable 4	Factor	Random factor with 50 levels

Subsamples of fish length measurements are available in the dataset, but this information could not be included in the random forest model because length sampling of catches is only done on a subset of fishing days with a maximum of 20 discarded and 20 landed individuals each week, meaning that data are only available for a small fraction of fishing operations. Including length as an additional variable must be done separately on the limited dataset and was therefore out of the scope of this study.

To mitigate bias in the random forest model arising from numeric variables on different scales, we standardised all numeric variables by subtracting the mean and dividing by two standard deviations (Gelman, 2008). We also included two continuous variables with randomly generated values from a normal distribution with mean = 0 and standard deviation = 1, and two random categorical variables with five and 50 levels, respectively. These four random variables will help to identify if the permutation method of evaluating variable importance is biased towards numeric variables, or categorical variables with differing number of levels (Ono *et al.*, 2016).

Random forests were fitted using the *ranger* package (Wright and Ziegler, 2017). The optimal random forest model was defined by the lowest number of variables randomly sampled at each node (*mtry*) and lowest number of trees *(ntree)* needed to minimise the mean square error of predictions. The tuning procedure resulted in the final model using parameters *mtry* = 4 and *ntree* = 5 000 (Figure S1; Supplementary Materials). To understand the uncertainty in variable importance, we estimated importance from 50 replicate random forest models fitted to the original dataset but with different random seeds. We explored the relationship between numbers discarded and explanatory variables using partial dependence plots (R package: *pdp* version 0.7.0; Greenwell, 2017), selecting specific interactions for visualisation depending on the outcome of the importance estimation.

Results

Discard estimates

In the coastal gillnet fisheries during the period 2012-2018, cod were discarded in half (49 %) of the observed fishing trips. Of those trips in which discarding occurred, 99 % of discarding events involved 20 individuals or fewer. An estimated number of 1 139 198 (95 % CI: 975 529 - 1 373 548) cod were discarded in the study period across the entire Norwegian coast with an average discard rate (Equation 2) of 0.55 % (95 % CI: 0.45-0.70 %). Whilst this discard rate is low, there are still important spatial and temporal variabilities to consider.

There was an overall weak decreasing trend in total numbers of discarded cod throughout the study period in both north and south management systems (Figure 2A). Trends in discard rates between north and south were dissimilar. North of 62 °N, where 88 % of all fishing activity occurred, discard rates averaged less than one cod per trip for most years (Figure 2B). On the other hand, estimated discard rates were an order of magnitude higher in southern areas and showed an overall increasing trend across years.

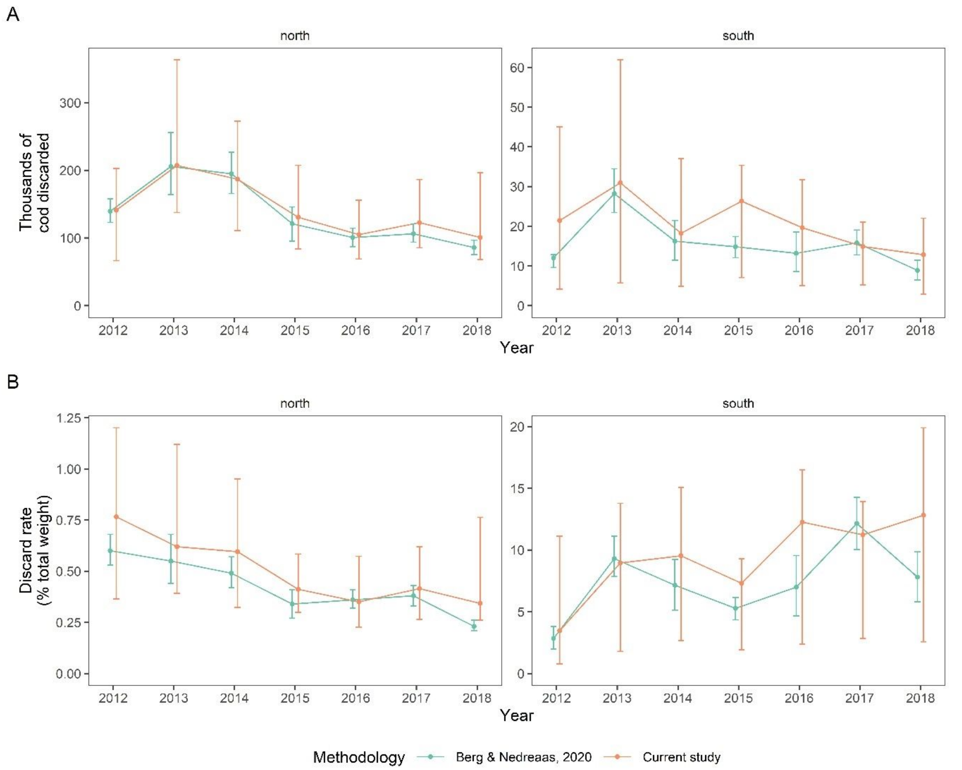


Figure 2. Annual estimated discards of cod in Norwegian coastal gillnet fisheries expressed as (A) total number discarded and (B) discard rate. Panels separate discards north and south of 62 °N (note different y-axes). Previous estimates by Berg & Nedreaas (2020) are included for comparison.

The revised methodology produces estimates with lower precision than previously reported by Berg & Nedreaas (2020). This increased uncertainty arises from accounting for variations in discarding across vessels and the conversion from total number of cod discarded (Figure 2A) to total weight to describe the discard rate (Figure 2B). The redefined assumptions for excluding fisheries and defining a fishing trip affected estimates in areas south of 62 °N much more than in the north. However, the trends remain very similar and previous estimates fall almost entirely into the range of uncertainty described in the revised methodology.

Accounting for additional sources of variance in the conversion from numbers to weight can result in large losses in precision (i.e., increase in CV; Figure 3). There are many strata for which ignoring the additional sources of variance results in an over-optimistic picture of precision. In these strata, including the additional variance causes CV to increase by as much as 70 %. However, if CV was already high before accounting for the additional variance, then their inclusion has negligible impacts on precision (increase in CV < 1 %).

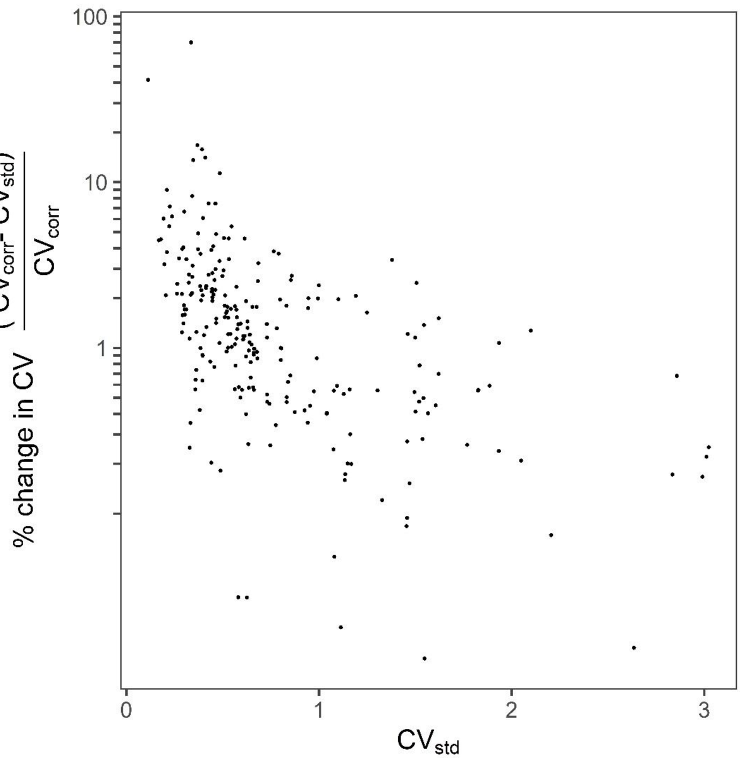


Figure 3. Percent change in coefficient of variation (CV) of discard rate estimates in individual strata (n = 252) when uncertainties in the numbers-to-weight conversion are included (CV_{corr}), compared to when they are ignored (CV_{std}). Strata are defined as year, statistical area (Figure 1) and calendar quarter.

The largest numbers of estimated discards were in the seasonal targeted spawning migration fishery on cod, which is confined to the Lofoten area of Norway (statistical areas 00, 04 and 05) in the first annual quarter (Figure 4A). However, when expressed as a discard rate (Equation 3; Figure 4B), values are low in these areas. In southern areas, particularly in Skagerrak and Kattegat (statistical area 09), and adjacent North Sea (statistical area 08) discard rates are higher.

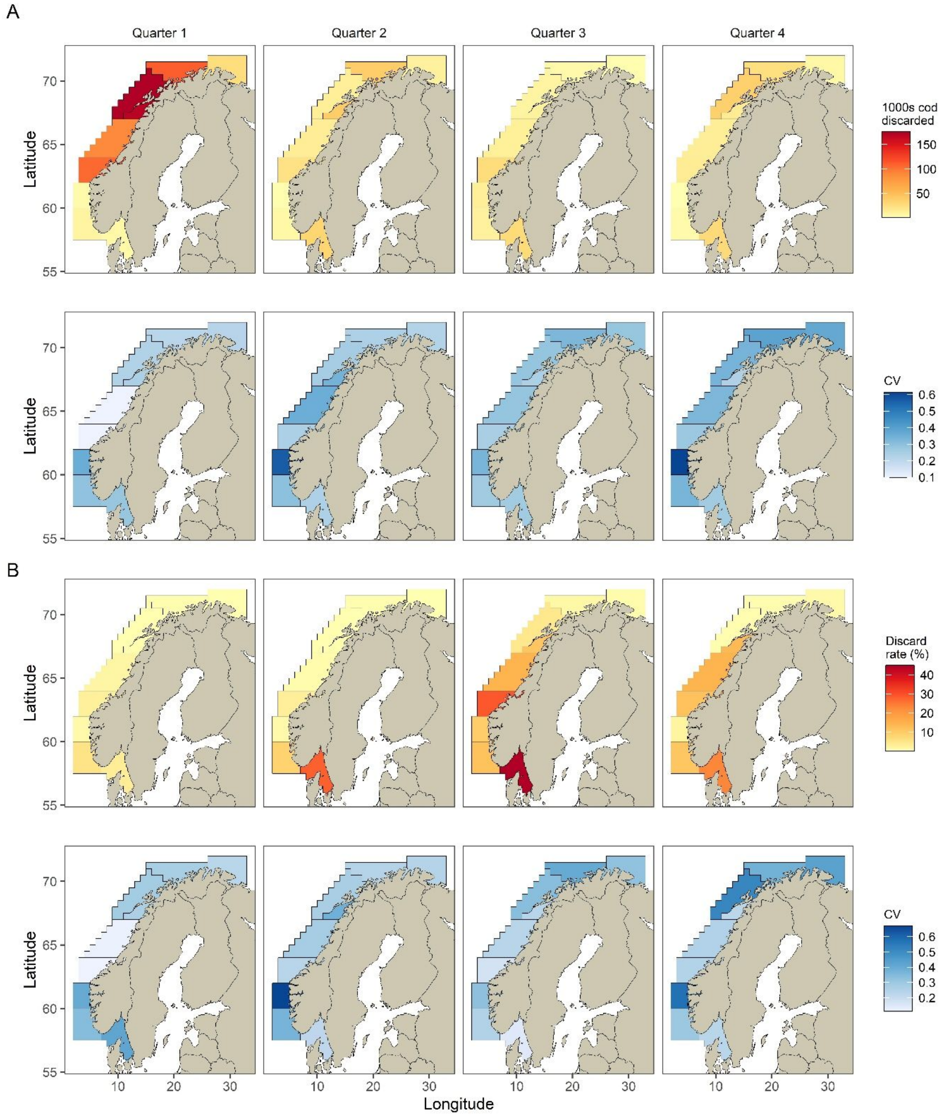


Figure 4. Spatial and temporal variations in discards of cod in Norwegian coastal gillnet fisheries expressed as (A) total numbers discarded and (B) discard rate, with associated uncertainty (coefficient of variation).

Drivers of discarding

The random forest model explained about half (44 %) of the variance in discarding of cod in the coastal gillnet fisheries using the 16 explanatory variables available. Of highest importance was the retained weight of cod in each fishing operation (Figure 5), alongside the soak time of nets. Fine-scale spatial variations were of higher importance than statistical area and management system which are on coarser scales. Of the variables explaining temporal variations, the random forest model found that price and month were most important, with annual and quarterly variations being relatively less important. However, some temporal trends may also be related to the landed weight of cod, which describes the large increases in catches of cod in the seasonal targeted cod fishery. Negligible importance of random variables suggested that these results can be interpreted without the risk of confounding biases from the permutation method.

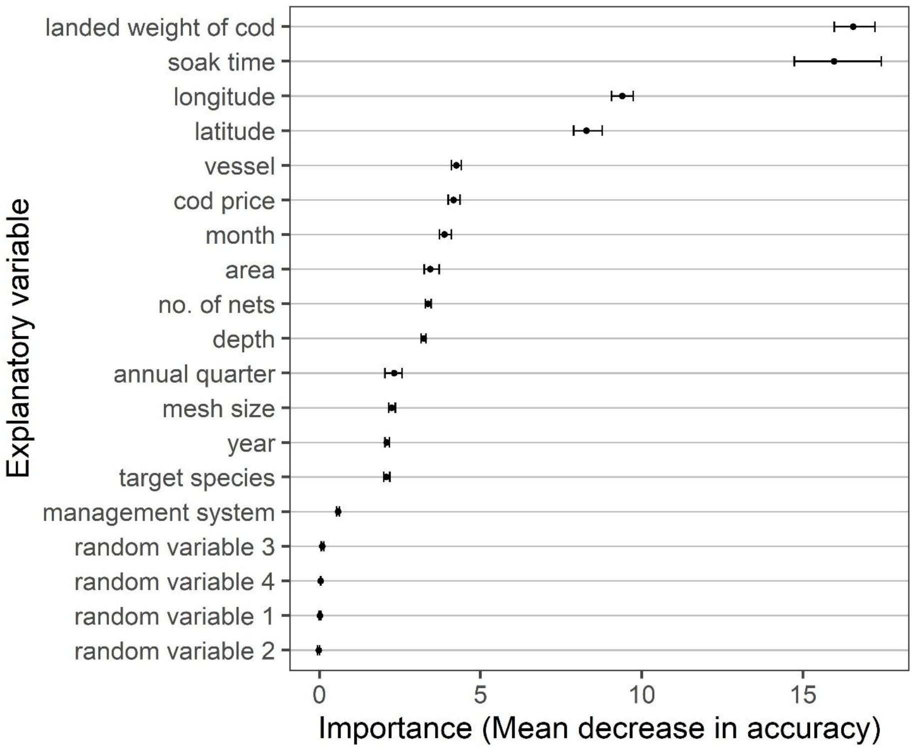


Figure 5. Importance of explanatory variables from a random forest model predicting discarded cod in the Norwegian coastal gillnet fisheries between 2012 and 2018. Importance is defined as the mean decrease in model accuracy when the values of each variable are randomly permuted. If a variable is important then model accuracy will decrease when values are randomly permuted because the association with discarding is destroyed. Estimates are mean and range of importance measures from 50 replicate random forests fitted to the original dataset with different random seeds.

When no cod were landed, an average of 1.2 (95 % CI: 0.0-4.6) cod were discarded per trip. As landed and reported catches of cod increased, discards increased proportionally, until reaching a saturation point at approximately 12 000 kg of landed cod, above which discards did not increase (Figure 6). A similar trend occurred with soak time of nets, where discarding did not increase above a soak time of ~100 hours. However, these interpretations should consider the reduced number of data points at the extreme values, particularly for soak time.

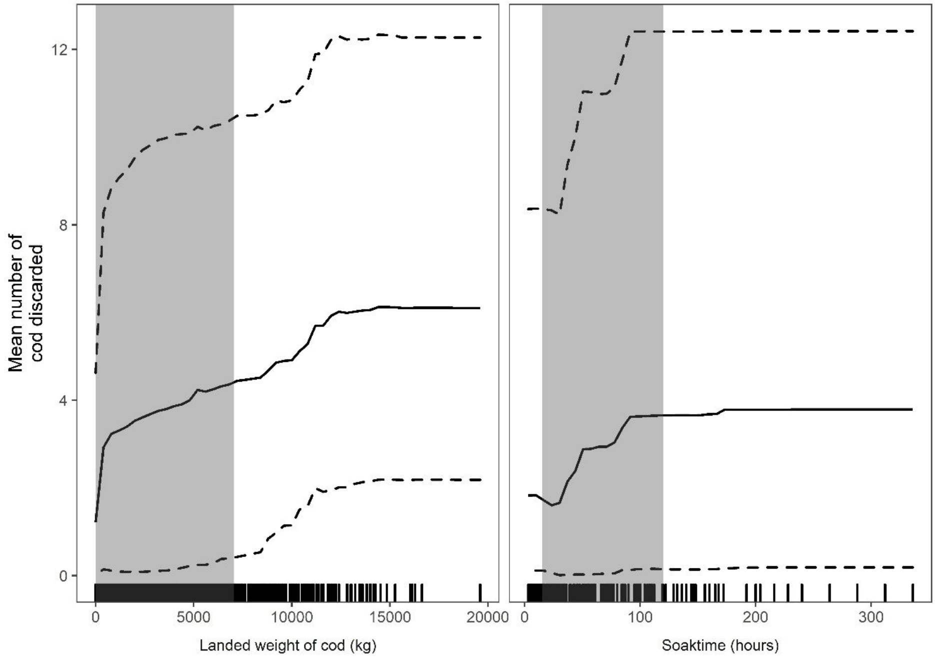


Figure 6. Partial dependence plots of selected important variables for predicting discards of cod in the Norwegian coastal gillnet fisheries between 2012 and 2018. Plots show the marginal effect of each variable on cod discards (solid line = mean; dashed lines = 95 % confidence interval). Tick marks along x-axes show the distribution of observations and the grey shaded area shows where 95 % of all observations lie.

Discarding increased as the price of cod decreased and quantities landed increased (Figure 7A). However, there are fewer observations with large quantities of landed cod, so caution is advised when interpreting trends. When there were no cod landed, discarding was not dependent of price, but was highly variable (Figure 7B). This variability can be explained by the variations in cod catches that were discarded and may indicate either over-quota discarding or unwanted catches in the non-target fisheries.

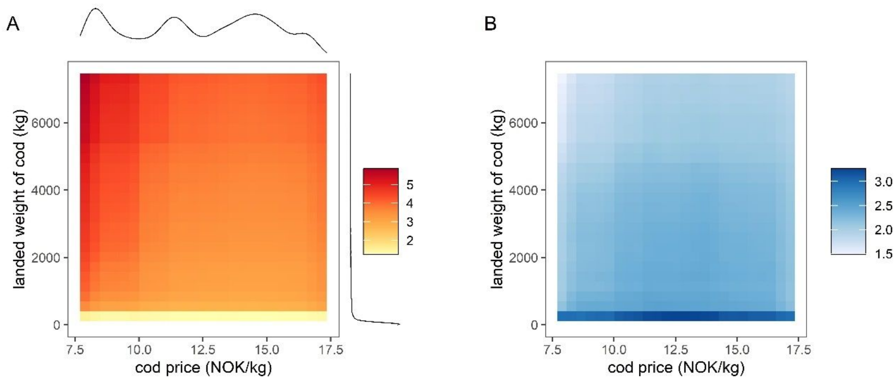


Figure 7. Interaction between cod price and landed weight of cod on estimated mean number of cod discarded per fishing day between 2012 and 2018. Data limited to 95 % range of observations. (A) Estimated number of discarded cod; (B)

uncertainty (standardised 95 % confidence interval). Marginal density plots in (A) show distribution of data for each of the explanatory variables.

Plotting fine-scale spatial variations (Figure 8) reveal that discarding is relatively homogenous within statistical areas, except for the Lofoten area (statistical areas 00, 04 and 05), where discarding is more variable (Figure 8A) and uncertain (Figure 8B). However, uncertainty is highest in mid Norway in statistical area 06 and 07.

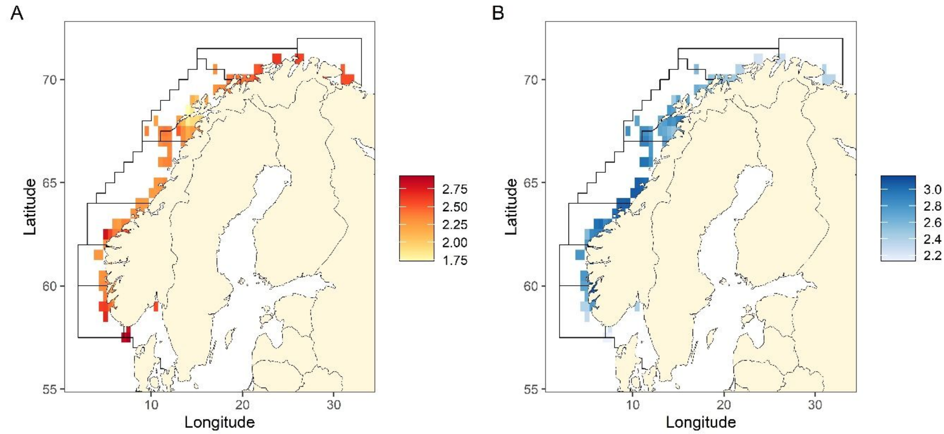


Figure 8. Influence of spatial variation on estimated mean number of cod discarded per fishing day between 2012 and 2018. (A) Estimated number; (B) uncertainty (standardised 95 % confidence interval). Partial dependence of latitude and longitude estimated for a 0.5 x 0.5 ° grid of observed fishing activity.

Discussion

Earlier estimates of discards in Norwegian fisheries (McBride and Fotland, 1996; Dingsør, 2001; Valdemarsen and Nakken, 2002; Nedreaas *et al.*, 2015) were based on inference or assumptions due to a lack of direct scientific sampling of discards. Berg and Nedreaas (2020) presented a generalised approach to estimating discards in the coastal gillnet fisheries using direct observations by the Norwegian Reference Fleet. This study developed the methodology further by accounting for both the clustered nature of sampling by the Norwegian Reference Fleet and additional uncertainties in the conversion from estimated numbers to weights. The stratification system is limited by the spatial information in sales notes, and the sampling effort limits a finer temporal scale of stratification. A ratio estimator based on soaking time of nets is unavailable, again due to a lack of information in sales notes, and using the landed weight of cod is unsuitable due to a non-linear relationship between landed and discarded cod (Figure 6; see also Rochet and Trenkel, 2005; Lohr, 2010). Using a different assumption to remove recreational and pelagic fisheries from the sampling frame has also resulted in a slight change in the annual trend. Whilst this change is quantifiable, any interpretations are masked by large, overlapping uncertainties. The exploratory modelling presented here has improved our understanding of the potential drivers of discarding. Strong fine-scale spatial variations and a dependence on fishing intensity (total catches of cod and soaking time) are the most important drivers of discarding, but we found that discarding was also explained by a complex combination of all other variables included in the model.

We included simple descriptors of fisheries (e.g., mesh size, target species) in the random forest model to suggest future improvements to the stratification of the design-based estimator. However, the model suggests that discarding variations across fisheries cannot be described by a simple categorisation. Instead, the model suggests that different variables may be characterising each fishery. For example, the landed weight of cod explains the degree to which cod is targeted in the mixed gadoid fishery, soak time helps to identify anglerfish nets that have longer soaking times, and an interaction of fine-scale spatial and temporal variables may pinpoint the targeted cod fishery that is isolated to the Lofoten area (statistical areas 00, 04, and 05) between January and April. These complex interactions should be an important consideration for future model-based estimators to ensure that all fisheries are well-described in the model.

Under the Norwegian discard ban, fishers are legally allowed to discard viable fish. It has also become a practice for the enforcement agencies not to prosecute discarding of damaged fish that are unfit for human consumption (Gullestad *et al.*, 2015). These exemptions must be considered when interpreting results, as discarding is used to correct catch data in stock assessments which assumes 100 % mortality. However, the survivability of discards in gillnet fisheries is dependent on a complex interaction of factors including species, gear specifications, soaking time, catch size and composition, air exposure and handling (Davis, 2002; Veldhuizen *et al.*, 2018, Sogn-Grundvåg *et al.*, 2022). Considering that coastal fisheries can have shorter soaking and handling times, the assumption of 100 % mortality is uncertain. Nevertheless, there are no survival studies to date that can be applied to discarded cod in coastal gillnet fisheries (ICES, 2020b), and the Norwegian Reference Fleet do not record the viability of discarded fish, or even the reason for discarding. However, even if 100 % discard mortality is assumed, the small estimated average discard rate of 0.55 % (95 % CI: 0.45-0.70 %) throughout the study period for this fishery will fall within the general uncertainties in stock assessments for cod and are likely negligible.

There were large differences in estimated discard rates between the areas north and south of 62 °N. These differences can likely be explained by the differences in regulations, fishing pattern and behaviour between the northern and southern parts of the Norwegian coast. In northern areas there is a much larger fishery for cod in general (88% of landed cod). This results in large total catches and in the targeted fishery for spawning northeast Arctic cod, catches almost exclusively consisting of larger, mature cod. This large targeting might explain why the estimated numbers of discarded cod are the highest in these areas, but with correspondingly low discard rates. In southern areas the cod stocks are smaller which results in cod being targeted to a lesser extent, and probably also consists of smaller individuals on average (Berg and Nedreaas, 2020). This results in higher estimated discard rates for cod, even if the estimated numbers of discarded cod are lower. Interpreting the results of discards in terms of consequences for the three cod stocks being caught in the coastal gillnet fishery is complex and due to data limitations, it has not been investigated further in this study.

This study found that discarding was driven by a complex combination of factors, most important of which were total catches of cod, soak time and fine-scale spatial variations. However, all other variables included, such as sampling units (vessels) and prices, were of importance to some degree. Whilst we have demonstrated the effectiveness of these data for describing variations in discarding, their usefulness for predicting discards in the wider fishery is limited by the lack of complementary data in the mandatory reporting system for all vessels. Historical estimates are limited in their stratification by the available data. For example, monthly variations in discarding were important (Figure 5), but there are too few observations to estimate monthly discard rates using a design-based approach. A similar issue occurs with fine scale spatial information, as the large scale of statistical areas were not very important for explaining variations if latitude and longitude were included in the model. The Norwegian Reference Fleet report the latitude and longitude of fishing operations, which allows for spatial modelling (Yan *et al.,* 2021), as well as attaching additional information that may explain discarding such as depth or habitat. However, the predictive capability of spatial modelling is limited if comparable data are not available for all fishing operations in the fishery.

Fish length is recognised as an important driver of discarding where minimum landing sizes are enforced (Batsleer *et al.*, 2005; Rochet and Trenkel, 2005; Borges *et al*., 2006). Even though discarding in coastal gillnet fisheries is very low on average, there is still a risk that discarding is size-based in terms of high-grading or undersized catches, as an inverse relationship between price and discards was found in this study (Figure 7). Ignoring the disproportionate impact of discarding on smaller fish could mask estimated recruitment trends in stock assessments (Punt *et al.*, 2006; Batsleer *et al.,* 2005). Our study could not address size-based discarding due to both the limited biological sampling of discards by the Norwegian Reference Fleet, and the added complications of the clustered structure of sampling (Nelson, 2014) that could not be accounted for sufficiently in our random forest model. However, graphical comparisons of length distributions of the discarded and retained catches of cod by Berg and Nedreaas (2020) have found variations in size-based discarding in both space and time. For example, size-based discarding was detected in the northernmost areas in the first two quarters when fishing intensity is highest, whilst in southern areas, discarding was relatively similar over all sizes. We therefore suggest a further study on size-based discarding of cod in the coastal gillnet fisheries, using the subset of discards data where length measurements were taken.

A new reporting system is gradually being rolled out in Norwegian coastal fisheries, requiring all vessels to report catches and fishing activity after every trip through an electronic reporting system alongside a mandatory vessel monitoring system for spatial tracking (Norwegian Ministry of Trade, Industry and Fisheries, 2009). Based on results from this study, this new reporting system will provide a wealth of data for improving fishery-scale predictions of total discards in coastal fisheries. We therefore suggest that model-based estimators can make use of these data to hopefully improve the precision and bias of estimates. However, we highlight that this study did not assess the predictive performance of the random forest model, which is important if this model will be used to estimate total discards in the fishery once the new reporting system is in full operation. The most important variables identified by the random forest model (Figure 5) can inform an information-theoretic approach to generalised linear model fitting (Burnham and Anderson, 2002), although correlated variables would need to be accounted for in this context (see Figure S2; Supplementary Materials).

Estimates presented here are based on data from a fishery self-sampling programme, and although the reference fleet is protected from prosecution and operating under a contractual agreement of accuracy, we must acknowledge the potential sources of bias. Participation in the Norwegian Reference Fleet is voluntary, and there is a possibility that the behaviour of the fishers on the participating vessels behave differently to the rest of the fishing fleet. Participation is paid, but the structure of this payment might influence the fishing behaviour of the participants. Reliability of self-sampled data must also be acknowledged. As discarding of viable fish is legal and data are not available to enforcement and control authorities, the risk that data are intentionally manipulated to avoid prosecution is negligible. This is supported by the fact that fishers are willing to report discards in the first place. However, there is also the risk of under-reporting discards if undesirable results could lead to a reduction in quotas or loss of fishing rights (Roman *et al.*, 2011). There is evidence that suggests the Norwegian Reference Fleet are equally willing to record discards as the wider fleet based on a study by Fangel *et al.* (2015) who found similar estimates of seabird bycatches in the coastal gillnet fisheries based on the Norwegian Reference Fleet and a questionnaire survey. We argue based on values that fishers willing to participate in the Norwegian Reference Fleet have an interest in the long-term sustainability of their fisheries and understand the impacts of data manipulation. There is always a risk that individual data collectors could manipulate data, even in independent observer programmes (Ewell *et al.*, 2020), but we believe that the issue is not systemic enough to incur substantial biases. Nevertheless, quantitative studies on reliability are needed, which would provide statistical evidence for these claims.

In conclusion, we have improved estimations of discards in the coastal gillnet fisheries for both past and future years, placing emphasis on the importance of accounting for uncertainties in each step of the analysis, and suggesting ways in which both accuracy and precision of future estimates can be improved when a higher resolution data collection programme is established. The methods are applicable to all other species in the gillnet fisheries and are robust in terms of precision, given that uncertainty is accounted for at every stage of sampling. The results presented here suggest that discarding is negligible. However, this cannot be concluded without determining how the corrected catch statistics will impact the stock assessment (Perretti *et al.*, 2020). Simulation studies have found that the impact of additionally fishing mortality incurred by discarding is also dependent on the trend in fishing effort (Dickey-Collas *et al*., 2007) and produce ‘unintuitive’ biases in estimates of stock status (Rudd and Branch, 2017). Furthermore, Berg and Nedreaas (2020) identified that discarding of cod in the coastal fisheries may be size-based, suggesting that some stock assessment parameters such as recruitment may be disproportionally affected by unknown levels of discarding.

Funding

This work was funded equally by the Norwegian Institute of Marine Research and the Norwegian Directorate of Fisheries.

Data availability statement

The data underlying this article cannot be shared publicly due the sensitivity of the contents and the privacy of fishers involved in data collection. The data will be shared on reasonable request to the corresponding author.

Author contributions

HSFB and TLC contributed equally to all aspects of the study. All other authors contributed to developing the methodology, interpretation of results, and editing of the manuscript.

Acknowledgements

The authors greatly appreciate the work by the Norwegian Reference Fleet in gathering and reporting data. Discussions and trips with the fishers have helped our understanding of the fishery. We also sincerely thank Tom Williams and Sofie Gundersen for their help with data and quality checking.

References

Aanes, S., and Pennington, M. 2003. On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. ICES Journal of Marine Science, 60: 297–303.

Batsleer, J., Hamon, K.G., van Overzee, H.M.J., Rijnsdorp, A.D., Poos, J.J. 2015. High-grading and over-quota discarding in mixed fisheries. Reviews in Fish Biology and Fisheries, 25: 715-736.

Benoît, H. P., and Allard, J. 2009. Can the data from at-sea observer surveys be used to make general inferences about catch composition and discards? Canadian Journal of Fisheries and Aquatic Sciences, 66: 2025–2039.

Berg, H.S.F., and Nedreaas, K. 2020. Estimering av utkast i norsk kystfiske med garn. Fisken og havet. Research report, 2021 - 1.

Borges, L., Zuur, A. F., Rogan, E., and Officer, R. 2006. Modelling discard ogives from Irish demersal fisheries. ICES Journal of Marine Science, 63: 1086–1095.

Boulesteix, A., Janitza, S., and Kruppa, J. 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. WIREs Data Mining Knowledge and Discovery, 2: 493–507.Breiman, L. 2001. Random forests. Machine Learning, 45: 5–32.

Breiman, L. 2001. Random forests. Machine Learning, 45: 5–32.

Burnham K.P., and Anderson D.R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag New York. 2nd edn.

Bærum, K.M., Anker-Nilssen, T., Christensen-Dalsgaard, S., Fangel, K., Williams, T., and Vølstad, J.H. 2019. Spatial and temporal variations in seabird bycatch: Incidental bycatch in the Norwegian coastal gillnet-fishery. PLOS ONE 14(3): e0212786.

Clegg, T., and Williams, T. 2020. Monitoring Bycatches in Norwegian Fisheries, Species registered by the Norwegian Reference Fleet 2015-2018. Rapport fra Havforskningen 2020-8

Clegg, T. L., Fuglebakk, E., Ono, K., Vølstad, J. H., and Nedreaas, K. 2021. A simulation approach to assessing bias in a fisheries self-sampling programme. ICES Journal of Marine Science, fsab242.

Cook, R. M. 2019. Inclusion of discards in stock assessment models. Fish and Fisheries: 1232–1245.

Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. J. 2007. Random forests for classification in ecology. Ecology, 88: 2783–2792.

Davis, M.W. 2002. Key principles for understanding fish bycatch discard mortality. Canadian Journal of Fisheries and Aquatic Sciences, 59(11): 1834–1843.

Depestele, J., Feekings, J., Reid, D. G., Cook, R., Gascuel, D., Girardin, R., Heath, M., *et al.* 2019. The Impact of Fisheries Discards on Scavengers in the Sea. *In* The European Landing Obligation: Reducing Discards in Complex, Multi-Species and Multi-Jurisdictional Fisheries, pp. 129–162. Springer, Cham.

Dickey-Collas, M., Pastoors, M. A., and van Keeken, O. A. 2007. Precisely wrong or vaguely right: simulations of noisy discard data and trends in fishing effort being included in the stock assessment of North Sea plaice. ICES Journal of Marine Science, 64: 1641–1649.

Dingsør, G. E. 2001. Estimation of discards in the commercial trawl fishery for Northeast Arctic cod (*Gadus morhua* L.) and some effects on assessment. Cand. Scient thesis, University of Bergen, 2001.
Efron, B., and Tibshirani, R. J. 1994. An Introduction to the Bootstrap. CRC Press. 456 pp.

Ewell, C., Hocevar, J., Mitchell, E., Snowden, S., and Jacquet, J. 2020. An evaluation of Regional Fisheries Management Organization at-sea compliance monitoring and observer programs. Marine Policy, 115: 103842.

Fangel, K., Aas, Ø., Vølstad, J.H., Bærum, K.M., Christensen-Dalsgaard, S., Nedreaas, K., Overvik, M., et al. 2015. Assessing incidental bycatch of seabirds in Norwegian coastal commercial fisheries: Empirical and methodological lessons. Global Ecology and Conservation, 4: 127–136.

Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. Statistics in Medicine, 27: 2865–2873.

Greenwell, B. M. 2017. pdp: An R Package for Constructing Partial Dependence Plots. The R Journal, 9: 421–436.

Gullestad, P., Blom, G., Bakke, G., and Bogstad, B. 2015. The "Discard Ban Package": Experiences in efforts to improve the exploitation patterns in Norwegian fisheries. Marine Policy, 54: 1–9.

Helle, K., and Pennington, M. 2004. Survey design considerations for estimating the length composition of the commercial catch of some deep-water species in the northeast Atlantic. Fisheries Research, 70: 55–60.

ICES, 2020a. Arctic Fisheries Working Goup (AFWG). ICES Scientific Reports, 2: 52. 577 pp.

ICES, 2020b. Working Group on Methods for Estimating Discard Survival (WGMEDS; outputs from 2019 meeting). ICES Scientific Reports, 2: 8. 75 pp.

ICES, 2021. Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK). ICES Scientific Reports. 3:66. 988 pp.

Karp, W. A., Breen, M., Borges, L., Fitzpatrick, M., Kennelly, S. J., Kolding, J., Nolde Nielsen, K., *et al.* 2019. Strategies Used Throughout the World to Manage Fisheries Discards – Lessons for Implementation of the EU Landing Obligation. *In* The European Landing Obligation: Reducing Discards in Complex, Multi-Species and Multi-Jurisdictional Fisheries, pp. 3–26. Ed. by S. S. Uhlmann, C. Ulrich, and S. J. Kennelly. Springer, Cham.

Kraan, M., Uhlmann, S., Steenbergen, J., van Helmond, A. T. M., and van Hoof, L. 2013. The optimal process of self-sampling in fisheries: lessons learned in the Netherlands. Journal of Fish Biology, 83: 963–973.

Lohr, S. L. 2010. Sampling: Design and Analysis. Chapman and Hall/CRC. 596 pp.

Mangi, S. C., Dolder, P. J., Catchpole, T. L., Rodmell, D., and de Rozarieux, N. 2013. Approaches to fully documented fisheries: practical issues and stakeholder perceptions. Fish and Fisheries, 16: 426–452.

McBride, M. M., and Fotland, Å. 1996. Estimation of unreported catch in a commercial trawl fishery. Journal of the Northwest Atlantic Fisheries Society, 18: 31–41.

Moan A., Skern-Mauritzen M., Vølstad J.H., and Bjørge A. 2020. Assessing the impact of fisheries-related mortality of harbour porpoise (*Phocoena phocoena*) caused by incidental bycatch in the dynamic Norwegian gillnet fisheries. ICES Journal of Marine Science, 77: 3039–3049.

Nedreaas, K., Iversen, S., and Grethe, K. 2015. Preliminary estimates of total removals by the Norwegian marine fisheries, 1950-2010. Working Paper #2015-94, Fisheries Centre, University of British Columbia, Vancouver, Canada.

Nelson, G. A. 2014. Cluster Sampling: A Pervasive, Yet Little Recognized Survey Design in Fisheries Research. Transactions of the American Fisheries Society, 143: 926–938.

Norwegian Ministry of Trade, Industry and Fisheries, 2009. Regulations on position reporting and electronic reporting for Norwegian fishing and catching vessels. Last amended 2021.01.18. FOR-2021-01-18-116.

Norwegian Ministry of Trade, Industry and Fisheries, 2014. Regulations on landing- and sales notes. Last amended 2020.02.27. FOR-2020-02-27-220.

Ono, K., Shelton, A. O., Ward, E. J., Thorson, J. T., Feist, B. E., and Hilborn, R. 2016. Space-time investigation of the effects of fishing on fish populations. Ecological Applications, 26: 392–406.

Pennington, M., and Helle, K. 2011. Evaluation of the design and efficiency of the Norwegian self-sampling purse-seine reference fleet. ICES Journal of Marine Science, 68: 1764–1768.

Pérez-Roda, A., Gilman, E., Huntington, T., Kennelly, S.J., Suuronen, P., Chaloupka, M., and Medley, P. 2019. A third assessment of global marine fisheries discards. FAO Fisheries Technical Aquaculture Paper no. 633. 78 pp. Rome, FAO.

Perretti C.T., Deroba J.J., and Legault C.M. 2020. Simulation testing methods for estimating misreported catch in a state-space stock assessment model. ICES Journal of Marine Science, 77: 911–920.

Punt, A. E., Smith, D. C., Tuck, G. N., and Methot, R. D. 2006. Including discard data in fisheries stock assessments: Two case studies from south-eastern Australia. Fisheries Research, 79: 239–250.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rochet, M.-J., and Trenkel, V. M. 2005. Factors for the variability of discards: assumptions and field evidence. Canadian Journal of Fisheries and Aquatic Sciences, 62: 224–235.

Roman, S., Jacobson, N., and Cadrin, S. X. 2011. Assessing the reliability of fisher self-sampling programs. North American Journal of Fisheries Management, 31: 165–175.

Rudd, M. B., and Branch, T. A. 2017. Does unreported catch lead to overfishing? Fish and Fisheries, 18(2): 313-323.

Rue, H., and Martino, S. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society, 71: 319–392.

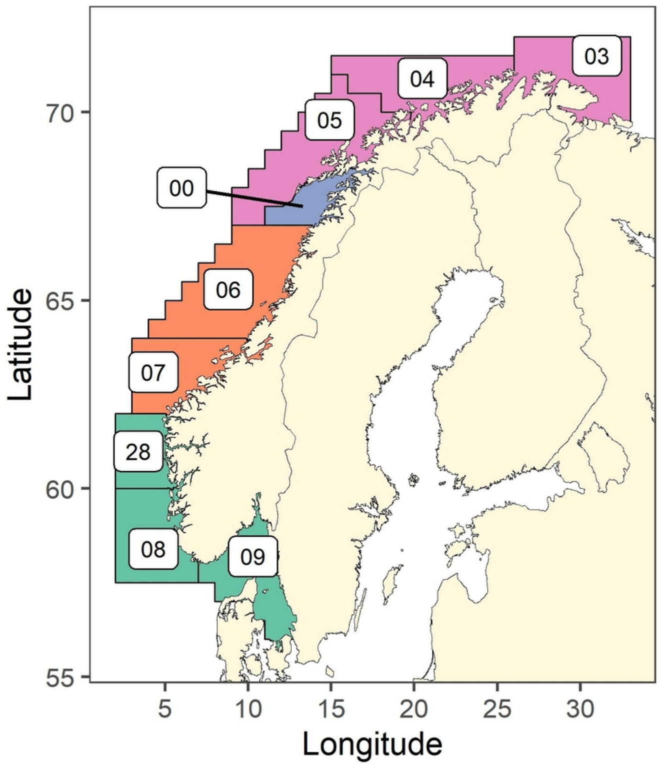
Sogn-Grundvåg, G., Zhang, D., Henriksen, E., Joensen, S., Bendiksen, B.-I., and Hermansen, Ø. 2022. Fishing tactics and fish quality: The case of the coastal fishery for Atlantic cod in Norway. Fisheries Research, 246: 106167.

Valdemarsen, J.W., and Nakken, O. 2002. Utkast i norske fiskerier, in: Workshop Om Utkast i Nordiske Fiskerier. Sophienberg Slot.

Veldhuizen, L. J. L., Berentsen, P. B. M., de Boer, I. J. M., van de Vis, J. W., and Bokkers, E. A. M. 2018. Fish welfare in capture fisheries: A review of injuries and mortality. Fisheries Research, 204: 41–48.

Wright, M. N., and Ziegler, A. 2017. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77: 1–17.

Yan, Y., Cantoni, E., Field, C., Treble, M., and Flemming, J. M. 2021. Spatiotemporal modeling of bycatch data: methods and a practical guide through a case study in a Canadian Arctic fishery. Canadian Journal of Fisheries and Aquatic Sciences: 1–11.



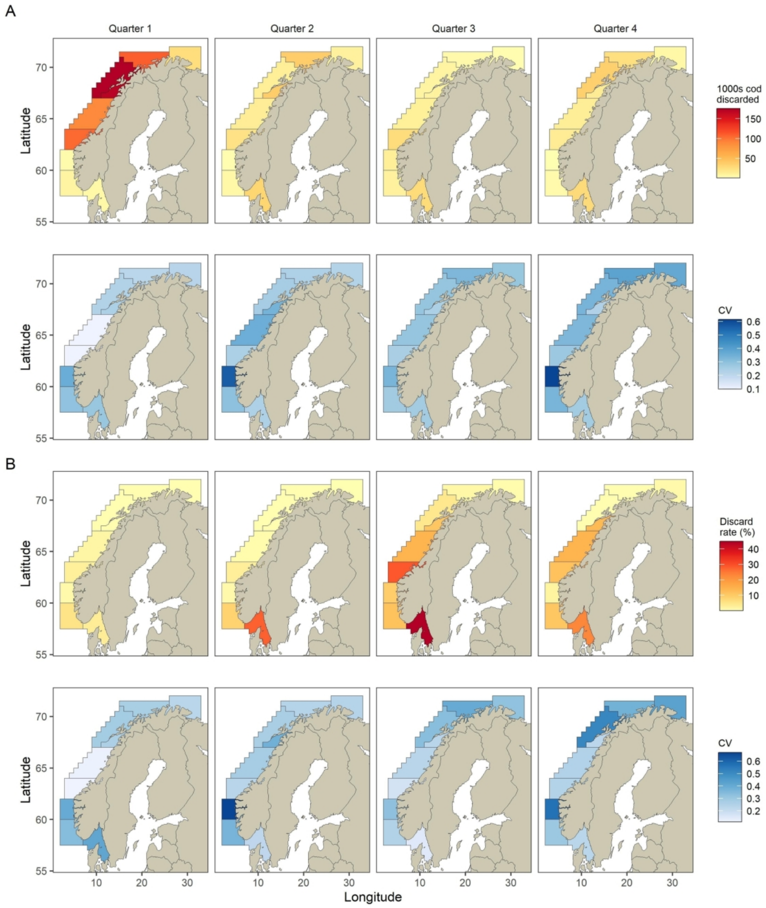
Map of study area including statistical areas defined by the Norwegian Directorate of Fisheries. Shading indicates regions used in estimation procedure. The division between north and south management systems is at 62 °N (the boundary between statistical areas 07 and 28). There was a small change in the geographic extent of areas 08 and 09 from 2018. This is accounted for in the analyses, but not shown in the maps presented.

84x84mm (300 x 300 DPI)

Unable to Convert Image

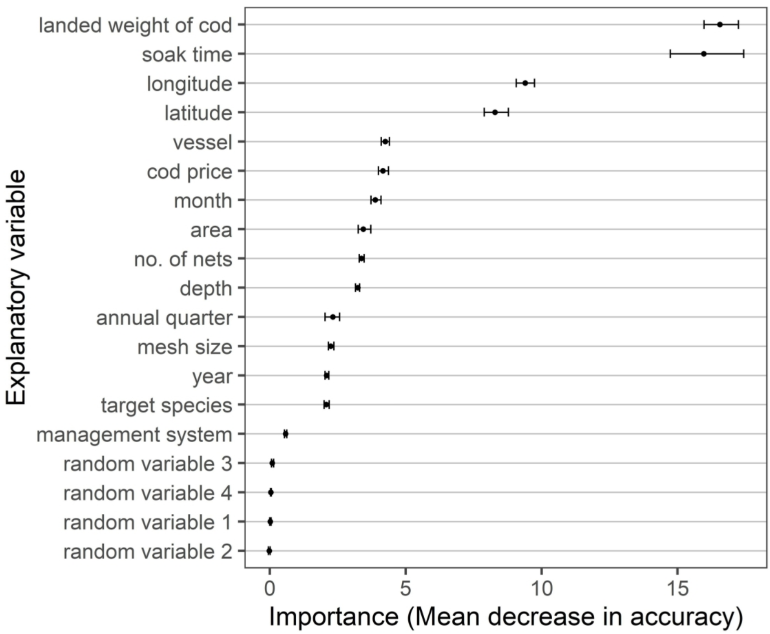
The dimensions of this image (in pixels) are too large to be converted. For this image to convert, the total number of pixels (height x width) must be less than 40,000,000 (40 megapixels).

Annual estimated discards of cod in Norwegian coastal gillnet fisheries expressed as (A) total number discarded and (B) discard rate. Panels separate discards north and south of 62 °N (note different y-axes). Previous estimates by Berg & Nedreaas (2020) are included for comparison.



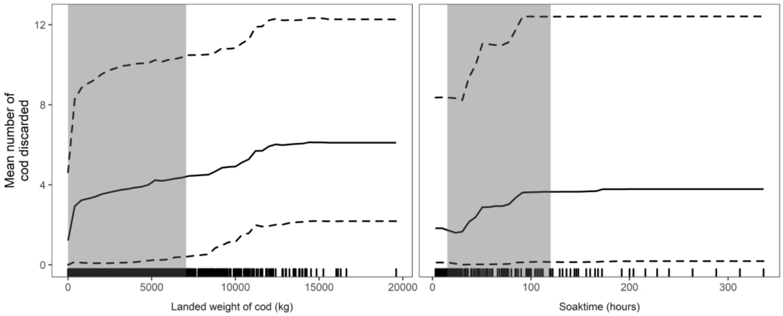
Spatial and temporal variations in discards of cod in Norwegian coastal gillnet fisheries expressed as (A) total numbers discarded and (B) discard rate, with associated uncertainty (coefficient of variation).

169x199mm (300 x 300 DPI)



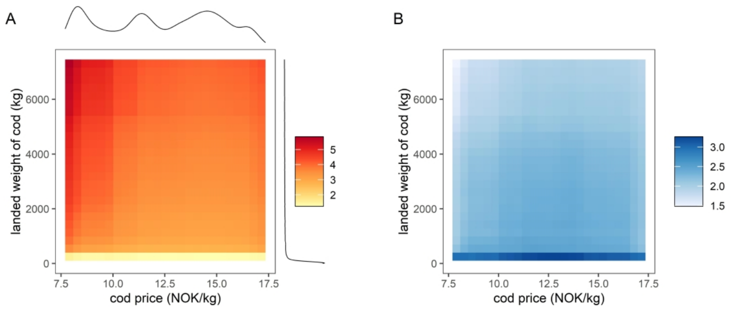
Importance of explanatory variables from a random forest model predicting discarded cod in the Norwegian coastal gillnet fisheries between 2012 and 2018. Importance is defined as the mean decrease in model accuracy when the values of each variable are randomly permuted. If a variable is important then model accuracy will decrease when values are randomly permuted because the association with discarding is destroyed. Estimates are mean and range of importance measures from 50 replicate random forests fitted to the original dataset with different random seeds.

84x69mm (600 x 600 DPI)



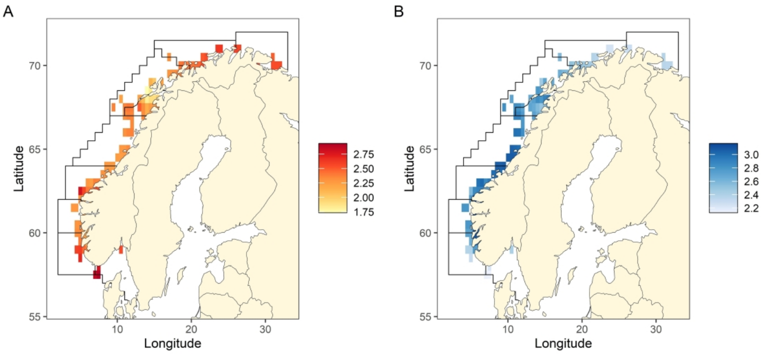
Partial dependence plots of selected important variables for predicting discards of cod in the Norwegian coastal gillnet fisheries between 2012 and 2018. Plots show the marginal effect of each variable on cod discards (solid line = mean; dashed lines = 95 % confidence interval). Tick marks along x-axes show the distribution of observations and the grey shaded area shows where 95 % of all observations lie.

169x69mm (600 x 600 DPI)



Interaction between cod price and landed weight of cod on estimated mean number of cod discarded per fishing day between 2012 and 2018. Data limited to 95 % range of observations. (A) Estimated number of discarded cod; (B) uncertainty (standardised 95 % confidence interval). Marginal density plots in (A) show distribution of data for each of the explanatory variables.

169x69mm (600 x 600 DPI)



Influence of spatial variation on estimated mean number of cod discarded per fishing day between 2012 and 2018. (A) Estimated number; (B) uncertainty (standardised 95 % confidence interval). Partial dependence of latitude and longitude estimated for a 0.5 x 0.5 ° grid of observed fishing activity.

169x89mm (600 x 600 DPI)

Supplementary Materials

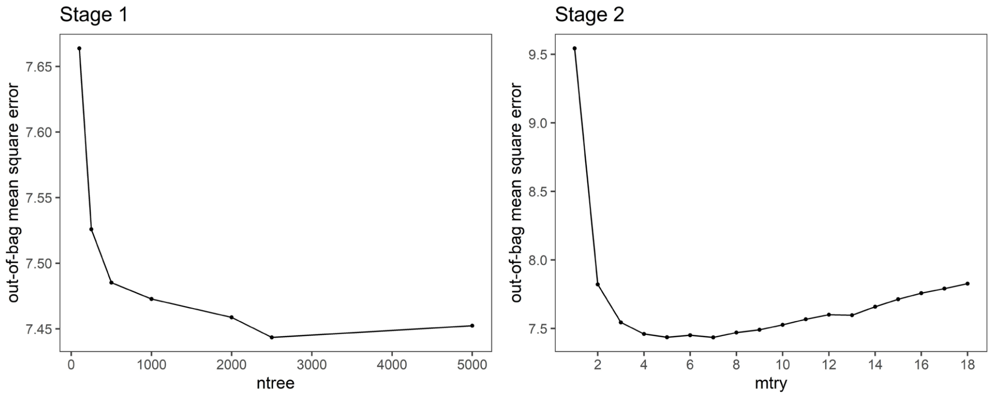


Figure S1. Tuning process of random forest. Stage 1 optimised the number of trees (ntree) whilst fixing the number of randomly sampled variables at each split (mtry) at the square root of the number of variables (mtry = 4). Stage 2 optimised mtry whilst using ntree = 5 000. Final model fitted using mtry = 4 and ntree = 5000.

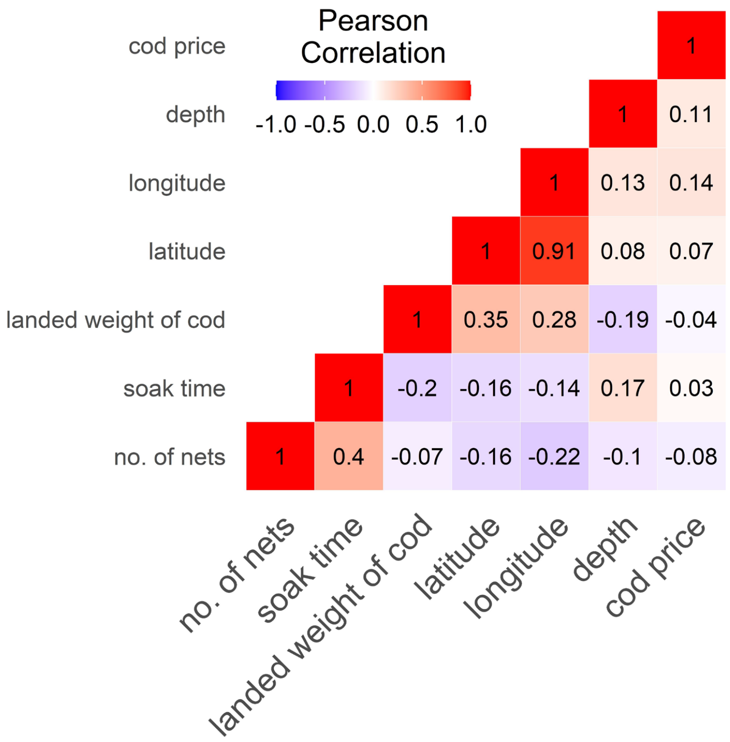


Figure S2. Correlation matrix of continuous variables used in the random forest model.

REPORT I

MONITORING BYCATCHES IN NORWEGIAN FISHERIES

Species registered by the Norwegian Reference Fleet 2015-2018

Tom Clegg and Tom Williams (IMR)

RAPPORT FRA
HAVFORSKNINGEN
NR. 2020-8

Title (English and Norwegian):

Monitoring bycatches in Norwegian fisheries

Overvåking av bifangst i Norske fiskerier

Subtitle (English and Norwegian):

Species registered by the Norwegian Reference Fleet 2015-2018

Arter registrerte av den Norske Referanseflåten 2015-2018

Report series:	**Year - No.:**	**Date:**
Rapport fra Havforskningen	2020-8	12.03.2020
ISSN:1893-4536		

Authors:

Tom Clegg and Tom Williams (IMR)

Godkjent av: Forskningsdirektør(er): Geir Huse Programleder(e): Elena Eriksen

Distribution:

Open

Project No.:

15561

Program:

Barentshavet og Polhavet

Research group(s):

Fiskeridynamikk

Number of pages:

26

Summary (English):

The Norwegian Reference Fleet is a group of active fishing vessels, selected as an approximate stratified random sample of vessels from the Norwegian fishing fleet, and tasked with providing information about catches and general fishing activity to the Institute of Marine Research. Fisheries data is collected by the crew members themselves, an approach commonly known as self-sampling of catches. This report aims to give an overview of how the Norwegian Reference Fleet record their catches and presents the reported catch composition with regards to number of species. A total of 271 species have been recorded by the Norwegian Reference Fleet between 2015 and 2018. There are an additional 39 records of unidentified species, which can occur because of excessive damage limiting an identification or a known misidentification that cannot be rectified.

Summary (Norwegian):

Referanseflåten er en gruppe aktive fiskefartøy, valgt ut som en tilnærmet stratifisert tilfeldig utvalg (stratified random sample) av fartøy fra den Norske fiskeflåten. Disse fiskefartøyenes hovedoppdrag for Havforskningsinstituttet er å bidra med informasjon om fangster og drift av fiskeriene. Fiskeridata er innsamlet ved såkalt «self sampling», hvor mannskapet om bord på fiskefartøyene selv utfører prøvetaking og dataregistrering. Formålet med denne rapporten er å redegjøre for hvordan Referanseflåten registrerer sine fangster og å presentere total fangstsammensetning i forhold til antall arter. Total har Referanseflåten registrert 271 arter mellom 2015 og 2018. I tillegg er det 39 registreringer av ikke identifiserte arter, som enten var ødelagte individer som ikke kunne identifiseres eller en bekreftet feilidentifisering som ikke kunne rettes.

Content

1 - Background and objectives

Monitoring bycatches in fisheries has become an integral part of fisheries management with regards to sustaining healthy ecosystems and the fisheries they support (Bellido et al. 2011). The Institute of Marine Research (IMR) in collaboration with the Norwegian fishing fleets, has developed the Norwegian Reference Fleet, a self-sampling programme used as a platform for supporting stock assessments with additional biological data including fishing effort, catch composition and bycatches. Since it was established in 2000, the data have been routinely used in stock assessments, but so far there have been relatively few publications on bycatch issues based on these data (e.g. Fangel et al. 2015; Bjørge & Moan 2017; Bærum et al., 2019). The aim of this report is to document the scope of sampling by the Norwegian Reference Fleet and provide an overview of the available data with regards to species reported in catches. A summary of species registered by the Norwegian Reference Fleet are provided in this report, along with the full dataset available for download (http://metadata.nmdc.no/metadata-api/landingpage/19d05ab8e0afe1ceac1b2be3ddf68612). Also included is an overview of the fisheries and fishing vessel categories that are prioritised in the Norwegian Reference Fleet, and the procedures used for reporting and sampling catches.

2 - The Norwegian Reference Fleet

2.1 – Aims of the project

The Norwegian Reference Fleet is a group of active fishing vessels tasked with providing information about catches and general fishing activity to the Institute of Marine Research. The fleet consists of both high-seas and coastal vessels that cover most of Norwegian waters. The High-seas Reference Fleet began in 2000 and was expanded to include coastal vessels in 2005. The four main goals of the Norwegian Reference Fleet are to:

1. Support stock assessments with biological data including:
   - Length composition of catches (length and weight measurements for all species captured)
   - Age composition of catches (otolith and scale collected)
   - Quality control and facilitation of data for stock-assessment
2. Document the fishing effort and catch composition of total catches, including bycatch, discards and catches of non-commercial species, seabirds and sea mammals to provide data for the monitoring of biodiversity, fishing effort and catch per unit effort (CPUE) over time
3. Provide a platform for the collection of additional samples from fisheries.
4. Increase collaboration and strengthen dialogue between researchers and the fishing industry.

2.2 – Vessel selection

The selection of vessels in the Norwegian Reference Fleet is required by law to follow an open tender process. The tender lists a series of criteria which are based on prioritised fisheries, vessel specifications and fishing gears (full description in Appendix Tables A1 and A2). These criteria prioritise data needed for stock assessments for commercially important stocks and reflect both spatial and temporal variation of fishing fleets. If multiple vessels are eligible under a certain category, then the contract is awarded randomly. The goal of the tender specifications and selection process is to approximate stratified random sampling, such that the Norwegian Reference Fleet is representative of the general fleet activity. A contract lasts for a period of four years, although renewal is possible if the vessel is still eligible.

For the larger vessels (>28m vessel length) in the Norwegian fishing fleet, the fisheries prioritised in the High-seas Reference Fleet are:

- demersal fisheries for cod, haddock and saithe north of latitude 62°N.
- demersal fisheries for cod, haddock and saithe south of latitude 62°N.
- beaked redfish trawl fishery.
- Greenland halibut fishery.
- ling and tusk fisheries with gillnet and longline.
- wolfish fishery with longline in the Barents Sea.
- pelagic fisheries with purse seine for herring, mackerel and saithe.
- industrial trawl fisheries south of latitude 62°N and in the North Sea targeting sandeel, Norwegian pout and blue whiting for fish-meal production.
- pelagic trawl fisheries for herring, mackerel, blue whiting and silver smelt.

For the smaller vessels (<28m vessel length) in the Norwegian fishing fleet, the fisheries prioritised in the Coastal Reference Fleet are:

- demersal fisheries for cod, haddock and saithe north and south of latitude 62°N (with particular focus on the Norwegian coastal cod component).
- Greenland halibut fishery.
- wrasse fishery with pots supplying cleaner fish to fish-farms.
- anglerfish fishery with gillnet.

- shrimp trawl fishery in the Skagerrak and North Sea.

In general, the demersal fisheries have been prioritised in both the High-seas and Coastal Reference Fleet, although for different reasons . Larger vessels in the demersal fisheries process their catches on board, meaning that at-sea sampling is necessary for obtaining length and age data of catches before they are processed. The fisheries prioritised in the Coastal Reference Fleet represent the most important fisheries in this sector of the Norwegian fishing fleet, which primarily target demersal species.

Vessels in the Norwegian Reference Fleet have the possibility to shift fisheries and target species, as long as it is in the constraints of the contract. This flexibility prevents excessive replacement of vessels due to vessels making small changes to their harvesting strategies, and because of the unpredictable nature of some fisheries. This means that there is a likelihood that not all prioritised fisheries will be covered by the Norwegian Reference Fleet each year. In addition, coastal fishing vessels are very adaptable to changes in the fisheries and can switch fishing gears and harvest strategies on very short timescales. Therefore, the Coastal Reference Fleet often provide additional data outside of the scope of the requirements and prioritised fisheries for each vessel category.

In 2019, the High-seas and Coastal Reference Fleet consisted of 16 and 22 vessels respectively (Appendix Tables A3 and A4). The number of vessels in the Norwegian Reference Fleet has been relatively stable throughout the period 2015–2019, with some vessels leaving the fleet after the contract period or for other reasons such as the fishing company selling the vessel. In each case, tenders were made to replace these vessels, although not always immediately after the contract was terminated.

2.3 - Sampling protocol and data handling

New vessels entering the Norwegian Reference Fleet are equipped with the necessary equipment and crew members are trained by IMR staff to ensure standardised sample processing and measurements. Alongside constant reporting of fishing activity and retained catches, bycatches and discards are also reported at regular intervals. The routine for documenting bycatches and discards in catches, and the sampling effort varies between fisheries and vessels (Appendix B). Bycatch of seabirds, sea mammals and rare fish species (e.g. porbeagle and basking shark) are also recorded for every fishing operation. From 2019, registering bycatch of corals and sponges is also included in the procedures.

Fishers are motivated to follow the protocol both through payment and an understanding of the importance of the collected data for stock assessment and management of the fisheries. Payment is effort based, with a price both for number of fish measured and number of species recorded in each catch, in order to give an incentive for fishers to use more time to follow the procedures correctly. The fishing vessels commitment to carry out this task is also outlined in the contract. There is an agreement between fishers, IMR and the relevant authorities that these data shall not be requested for enforcement purposes. This ensures that vessels can honestly report their catches without risk of prosecution, ensuring the data reflects the true catches. It is important to note that to date, this agreement has not been compromised.

Data are recorded electronically and regularly delivered to a database at IMR, where assigned IMR staff run quality control checks before approval. IMR staff are in regular contact with crew and skippers, and visit the vessels to provide support for self-sampling. C rew are also given training on species identification and new equipment both at sea and on land, and are issued the necessary literature to assist in species identification. If crew are uncertain about a species, they are encouraged to send photographs or samples to IMR for verification by taxonomists.

3 - Species registered by the Norwegian Reference Fleet

Data from Norwegian Reference Fleet vessels targeting Norwegian fish stocks between 2015 and 2018 is shown in Figure 1. Data from 2019 were incomplete at the time of publication and are therefore not included in this report. Species lists were generated for fishing gears used by the High-seas and Coastal Reference Fleet, divided between two areas north and south of 62°N latitude. Not all fishes were identified to species level, and are therefore grouped separately, whilst animals in other species groups were identified to different taxonomic levels.

A comprehensive list of registered species has been archived by the Norwegian Marine Data Centre at IMR (http://metadata.nmdc.no/metadata-api/landingpage/19d05ab8e0afe1ceac1b2be3ddf68612), and is summarised by species group in Figure 2. Tables 1-4 list the 30 most common species registered by vessel category. For each fishing gear, Table 2 lists the fisheries represented by target species. A total of 271 species have been recorded in 33,381 fishing operations by the Norwegian Reference Fleet between 2015 and 2018. There are an additional 39 records of unidentified species, which occur from issues flagged during quality control that cannot be rectified.

The list includes both landed and discarded species, but it is important to note that the Norwegian Reference Fleet do not record whether an animal was dead or alive when discarded. Reported quantities of catches are not provided as they are based on the relevant sampling protocols for a fishing gear. Therefore, reliable estimates of total catches for any given species in a fishery require dedicated methods for extrapolation, which is out of the scope of this report.

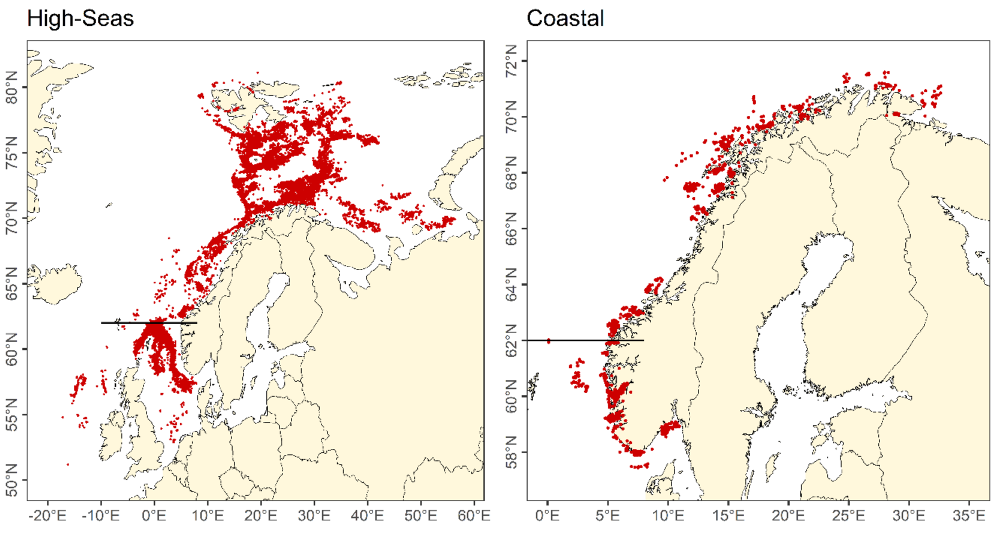


Figure 1 Locations of samples taken by the High-Seas and Coastal Reference Fleet between 2015 and 2018. Black horizontal line is at 62 °N latitude showing the division of north and south areas.

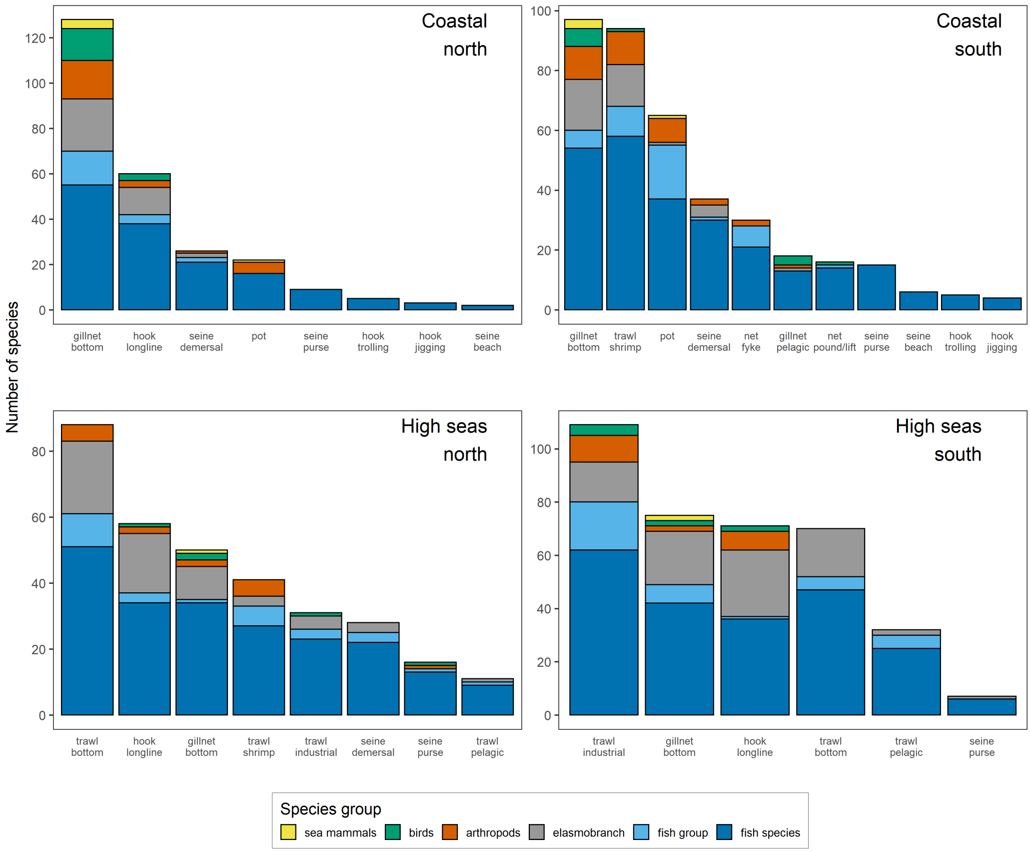


Figure 2. Summary of species registered by the Norwegian Reference Fleet. North/south is relative to 62°N latitude.

Table 1. List of the most common species registered tin total catches by the High-seas Reference Fleet, **north** of 62°N latitude. Species are listed in descending order with the most regular occurring species in the top row.

Gillnet bottom-set	Hook longline	Seine demersal	Seine purse	Trawl bottom	Trawl industrial	Trawl pelagic	Trawl shrimp
Atlantic cod	Atlantic cod	Atlantic cod	Saithe	Atlantic cod	Blue whiting	Saithe	Deep sea shrimp
Saithe	Haddock	Saithe	Atlantic herring	Haddock	Greater argentine	Atlantic herring	Long rough dab
Haddock	Starry skate	Haddock	Atlantic cod	Golden redfish	Saithe	Redfishes	Deepwater redfish
Ling	Spotted catfish	Ling	Haddock	Saithe	Atlantic herring	Blue whiting	Capelin
Golden redfish	Northern wolffish	Tusk	Mackerel	Deepwater redfish	Redfishes	Greater argentine	Polar cod
Tusk	Long rough dab	Atlantic halibut	Capelin	Starry skate	Haddock	Spurdog	Sclerocrangon
Pollack	Tusk	Golden redfish	Bluefin tuna	Greenland halibut	Argentines	Atlantic cod	Spotted snake blenny
Long rough dab	Atlantic catfish	Atlantic catfish	Gulls	Spotted catfish	Mackerel	Haddock	Atlantic hookear sculpin
Atlantic halibut	Golden redfish	Anglerfish (monk)	Tusk	Long rough dab	Golden redfish		Snakeblenny
Greenland halibut	Greenland halibut	Lumpsucker	Anglerfish (monk)	Atlantic catfish	Lanternfishes		Atlantic cod
Rabbitfish	Round skate	European plaice	Atlantic halibut	Lumpsucker	Porbeagle shark		Atlantic poacher
Blackmouthed dogfish	Atlantic halibut	Long rough dab	Blue whiting	Northern wolffish	Velvet belly		Lycodes
Starry skate	Ling	Redfishes	Ling	Atlantic halibut	European hake		Sea tadpole
European hake	Saithe	Greater argentine	Lumpsucker	Tusk	Ling		Greenland halibut
Atlantic herring	Deepwater redfish	European hake	Red king crab	Flounder	Silvery pout		Snailfishes
Anglerfish (monk)	Rough rattail	Lemon sole	Salmons	Greater argentine	Anglerfish (monk)		Shrimps
Spurdog	Spinytail skate	Spotted catfish		Ling	Atlantic cod		Haddock
Whiting	Rabbitfish	Whiting		Blue whiting	Blackmouthed dogfish		Prawns
European plaice	Greater forkbeard	Deepwater redfish		Round skate	Dealfish		Spotted catfish
Greater forkbeard	Esmark's eelpout	Flatfishes		Whiting	Deepwater redfish		Threespot eelpout
Spotted catfish	Blackmouthed dogfish	Starry skate		Spinytail skate	Greater forkbeard		White barracudina
Deepwater redfish	Arctic skate	Rabbitfish		Norway redfish	Long rough dab		Eelpouts
Northern wolffish	Velvet belly	Grey gurnard		Greater forkbeard	Norway pout		Glacial eelpout

Gillnet bottom-set	Hook longline	Seine demersal	Seine purse	Trawl bottom	Trawl industrial	Trawl pelagic	Trawl shrimp
Megrim	Norway redfish	Righteye flounders		Rabbitfish	Norway redfish		Snow crab
Rough rattail	Blue skate	Greater forkbeard		Lemon sole	Pollack		Golden redfish
Atlantic catfish	European plaice	Skates and rayes		Pollack	Spurdog		Starry skate
Norway redfish	Blue ling	Spurdog		Megrim	Whiting		Atlantic catfish
Lumpsucker	Roundnose grenadier			Anglerfish (monk)			Barracudinas
Redfishes	Spurdog			Esmark's eelpout			Bigeye sculpin
Round skate	Whiting			European hake			Shorthorn sculpin

Table 2. List of the most common species registered tin total catches by the High-seas Reference Fleet, **south** of 62°N latitude. Species are listed in descending order with the most regular occurring species in the top row.

Gillnet bottom-set	Hook longline	Seine purse	Trawl bottom	Trawl industrial	Trawl pelagic
Atlantic cod	Ling	Atlantic herring	Saithe	Blue whiting	Blue whiting
Saithe	Haddock	Mackerel	Ling	Norway pout	Mackerel
Haddock	Atlantic cod	Saithe	European hake	Saithe	Norway pout
Ling	Tusk	Atlantic cod	Atlantic cod	European hake	Atlantic herring
European hake	Saithe	Grey gurnard	Haddock	Silvery pout	Horse mackerel
Anglerfish (monk)	Small-spotted catshark		Mackerel	Atlantic cod	Argentines
Whiting	Cuckoo ray		Grey gurnard	Ling	Saithe
Pollack	Blue skate		Anglerfish (monk)	Argentines	European hake
Mackerel	Whiting		Tusk	Horse mackerel	Silvery pout
Starry skate	Pollack		Megrim	Anglerfish (monk)	Ling
European plaice	Spurdog		Atlantic herring	Haddock	Anglerfish (monk)
Tusk	European hake		Lemon sole	Witch	Atlantic cod
Spurdog	Anglerfish (monk)		Horse mackerel	Argentine	Whiting
Small-spotted catshark	Atlantic catfish		Blue whiting	Mackerel	Long rough dab
Witch	Starry skate		Greater argentine	Velvet belly	Argentine
Atlantic halibut	European conger eel		Starry skate	Whiting	Haddock
Megrim	Grey gurnard		Pollack	Atlantic herring	Pollack
Horse mackerel	Blackmouthed dogfish		Whiting	Long rough dab	Velvet belly
Atlantic catfish	Greater forkbeard		Atlantic halibut	Pollack	Hakes
Grey gurnard	Shagreen ray		Cuckoo ray	Pearlside	Atlantic catfish
Long rough dab	Triglops		Triglops	Blackmouthed dogfish	Boarfish
Tub gurnard	Rabbitfish		Witch	Blue-mouth redfish	Greater argentine
Atlantic herring	Longnosed skate		Greenland halibut	Spurdog	Rockfishes
Cuckoo ray	Atlantic halibut		Deepwater redfish	Poor cod	Triglops
Longnosed skate	Thornback ray		Greater forkbeard	Sand eel	Witch
Lemon sole	Blue-mouth redfish		Atlantic catfish	Atlantic catfish	
Spotted ray	Deepwater redfish		Blackmouthed dogfish	Tusk	
Turbot	European plaice		Golden redfish	Grey gurnard	
Dab	Golden redfish		Long rough dab	Greater forkbeard	
Starry smooth-hound	Sandy ray		Roundnose grenadier	Norway lobster	

Table 3. List of the most common species registered tin total catches by the Coastal Reference Fleet, **north** of 62°N latitude. Species are listed in descending order with the most regular occurring species in the top row.

Gillnet bottom-set	Hook longline	Other	Pot	Seine demersal	Seine purse
Edible crab	Haddock	Mackerel	Edible crab	Atlantic cod	Atlantic herring
Atlantic cod	Saithe	Saithe	Tusk	Haddock	Mackerel
Stone crab	Atlantic cod	Pollack	European plaice	Saithe	Saithe
Saithe	Tusk	Atlantic herring	Atlantic cod	European plaice	Atlantic cod
Haddock	Golden redfish	Horse mackerel	Red king crab	Anglerfish (monk)	Pollack
Ling	Atlantic halibut	Atlantic cod	European lobster	Lumpsucker	Horse mackerel
Atlantic halibut	Ling	Whiting	Atlantic catfish	Atlantic halibut	European hake
Pollack	Whiting		European conger eel	Megrim	Haddock
Anglerfish (monk)	Velvet belly		Shorthorn sculpin	Atlantic catfish	Whiting
Tusk	Blackmouthed dogfish		Common harbour seal	Ling	
Rabbitfish	Mackerel		Saithe	Dab	
Golden redfish	Norway redfish		Norway lobster	Norway pout	
European hake	Rabbitfish		Atlantic halibut	Spotted catfish	
Megrim	Atlantic catfish		Common dragonet	Turbot	
European plaice	Greenland halibut		Fourbeard rockling	Tusk	
Lemon sole	Skates and rayes		Hooknose	Grey gurnard	
Whiting	Grey gurnard		Ling	Pollack	
Blackmouthed dogfish	Starry skate		Shore rockling	Whiting	
Norway redfish	Pollack		Stone crab	Redfishes	
Starry skate	Greater forkbeard			Brill	
Lumpsucker	European hake			Golden redfish	
Spurdog	Spotted catfish			Lemon sole	
Grey gurnard	Deepwater redfish			Thornback ray	
Poor cod	Anglerfish (monk)			Norway lobster	
Velvet belly	Spurdog			Rockfishes	
Thornback ray	Redfishes			Spotted ray	
Greater forkbeard	Rough rattail				
Small-spotted catshark	Horse mackerel				
Mackerel	European plaice				
Atlantic herring	Edible crab				

Table 4. List of the most common species registered tin total catches by the Coastal Reference Fleet, **south** of 62°N latitude. Species are listed in descending order with the most regular occurring species in the top row.

Gillnet bottom-set	Gillnet pelagic	Net fyke	Other	Pot	Seine demersal
Stone crab	Mackerel	Atlantic cod	Mackerel	Corkwing	Atlantic cod
Atlantic cod	Atlantic herring	Ballan wrasse	Horse mackerel	Goldsinny wrasse	Haddock
Pollack	Saithe	Corkwing	Pollack	Ballan wrasse	European plaice
Ling	Garfish	Cuckoo wrasse	Saithe	Cuckoo wrasse	Anglerfish (monk)
Rabbitfish	Lumpsucker	Goldsinny wrasse	Greater sand eel	Edible crab	Pollack
Edible crab	Pollack	Pollack	Atlantic herring	Smallmouthed wrasse	Grey gurnard
Saithe	Spurdog	Poor cod	Atlantic salmon	European eel	Dab
Haddock	European hake	Smallmouthed wrasse	Whiting	Green shore crab	Turbot
Anglerfish (monk)	Razorbill	Bullheads and sculpins	Atlantic cod	Atlantic cod	Atlantic halibut
European hake	Trout	Green shore crab	Grey gurnard	European lobster	Saithe
Spurdog	Atlantic cod	Yarrell's blenny	Garfish	Bullheads and sculpins	Lemon sole
Velvet belly	Atlantic salmon	European eel	Red mullet	Pollack	Spurdog
Megrim	Ballan wrasse	Black goby	Sand lances	Poor cod	Brill
Norway redfish	Common eider	Edible crab	Blue whiting	Saithe	Ling
Tusk	Cuckoo wrasse	Viviporous eelpout	Cormorants	Tadpole fish	Megrim
Witch	Edible crab	Shanny	Poor cod	Shanny	Whiting
Blackmouthed dogfish	Northern fulmar	Ling	Rainbow trout	Black goby	European hake
Grey gurnard	Whiting	Saithe		Viviporous eelpout	John dory
Poor cod		Lemon sole		Ling	Skates and rayes
Lemon sole		Righteye flounders		Fivebeard rockling	Thornback ray
Blue ling		Common topknot		Gobies	Tub gurnard
Blue whiting		Eels		Munida	Atlantic catfish
Starry skate		Whiting		Butterfish	Flounder
Horse mackerel		Zoarcoids		Hyas	Greater weever
Atlantic halibut		Butterfish		Yarrell's blenny	Lumpsucker
Long rough dab		Flatfishes		Three-bearded rockling	Rabbitfish
Mackerel		Goatfishes		Shorthorn sculpin	Righteye flounders
Turbot		Pricklebacks		Common topknot	Stone crab
Whiting		Tadpole fish		Rocklings	Witch
Longnosed skate		Trout		Norway bullhead	Edible crab

Table 5. Description of target species for each fishing gear used by the Norwegian Reference Fleet. Area is relative to 62°N latitude.

Gear type	Area	Fleet	Vessel categories	Target Species
Hook jigging	North	Coastal	Gillnet/longline vessels north Gillnet/longline vessel south	Cod, saithe
	South	Coastal	Gillnet/longline vessel south	Cod, saithe, pollock, mackerel
Hook longline	North	High-seas	Longline/gillnet vessel	Cod, haddock, saithe, wolffish, ling, tusk, Greenland halibut
	South	High-seas	Longline/gillnet vessel	Cod, haddock, saithe, ling, tusk
	North	Coastal	Gillnet/longline vessels north	Cod, haddock, saithe, ling, tusk, Greenland halibut
Hook trolling	North	Coastal	Gillnet/longline vessel south	Mackerel
	South	Coastal	Gillnet/longline vessel south	Mackerel
Gillnet bottom-set	North	High-seas	Longline/gillnet vessel Gillnet vessel (Barents Sea)	Cod, haddock, saithe, ling, tusk, Greenland halibut
			Gillnet vessel (North Sea)	Cod
	South	High-seas	Longline/gillnet vessel Gillnet vessel (North Sea)	Cod, haddock, saithe, ling, tusk
	North	Coastal	Gillnet/longline vessels north.	Cod, haddock, saithe, ling, tusk, Greenland halibut, anglerfish
			Gillnet/longline vessel south Shrimp trawler (9-15m)	Cod, haddock, saithe
	South	Coastal	Gillnet/longline vessel south	Cod, haddock, saithe, ling, tusk
			Shrimp trawler (9-15m)	Cod, haddock, saithe
Gillnet pelagic	South	Coastal	Gillnet/longline vessel south	Mackerel
Net fyke	South	Coastal	Gillnet/longline vessel south	Wrasse, cod??
Net pound/lift	South	Coastal	Gillnet/longline vessel south	Mackerel
Pot	North	Coastal	Demersal seine vessel south	Mackerel
			Gillnet/longline vessel south	Wrasse, brown crab, *Nephrops*
	South	Coastal	Gillnet/longline vessel south	Wrasse, brown crab, *Nephrops*
			Demersal seine vessel south	*Nephrops*
Seine demersal	North	High-seas	Demersal/purse seine vessel	Cod, haddock
	North	Coastal	Demersal seine vessel north Demersal seine vessel south	Cod, haddock, saithe
	South	Coastal	Demersal seine vessel south	Cod, haddock, saithe
			Shrimp trawler 9-15m	Cod
Seine purse	North	High-seas	Demersal/purse seine vessel	Saithe, herring, mackerel, sprat, horse mackerel
			Industry trawler	Herring
	North	Coastal	Gillnet/longline vessels north	Herring
			Demersal seine vessel south Gillnet/longline vessel south	Herring, mackerel
	South	Coastal	Demersal seine vessel south	Mackerel, horse mackerel
			Gillnet/longline vessel south	Mackerel

Gear type	Area	Fleet	Vessel categories	Target Species
Seine beach	North	Coastal	Gillnet/longline vessels north.	Herring
	South	Coastal	Gillnet/longline vessel south	Mackerel
Trawl demersal	North	High-seas	Demersal factory trawler	Cod, haddock, saithe, Greenland halibut, beaked redfish
	South	High-seas	Demersal factory trawler	Saithe, Greenland halibut
Trawl industrial	North	High-seas	Industry trawler	Blue whiting, silver smelt, saithe
	South	High-seas	Industry trawler	Sandeel, Norwegian pout, blue whiting, saithe
Trawl pelagic	North	High-seas	Demersal factory trawler	Beaked redfish
			Industry trawler	Herring, mackerel
	South	High-seas	Industry trawler	Herring, mackerel, blue whiting, sprat
Trawl shrimp	North	High-seas	Demersal factory trawler	Shrimp
	South	Coastal	Shrimp trawler 9-15m Shrimp trawler 15-28m	Shrimp

4 - References

Bellido JM, Santos MB, Pennino MG, et al (2011) Fishery discards and bycatch: solutions for an ecosystem approach to fisheries management? Hydrobiologia 670:317–333. https://doi.org/10.1007/s10750-011-0721-5

Bjørge A, Moan A (2017) Revised estimates of harbour porpoise (Phocoena phocoena) bycatches in two Norwegian coastal gillnet fisheries. ICES By-catch WG. SC/24/BYCWG/08

Bærum KM, Anker-Nilssen T, Christensen-Dalsgaard S, et al (2019) Spatial and temporal variations in seabird bycatch: Incidental bycatch in the Norwegian coastal gillnet-fishery. PLoS One 14. https://doi.org/10.1371/journal.pone.0212786

Fangel K, Aas Ø, Vølstad JH, et al (2015) Assessing incidental bycatch of seabirds in Norwegian coastal commercial fisheries: Empirical and methodological lessons. Glob Ecol Conserv 4:127–136. https://doi.org/10.1016/j.gecco.2015.06.001

5 - Appendices

5.1 – Appendix A: General information on the Norwegian Reference Fleet

Table A1. Vessel requirements in the High-Seas Reference Fleet

Category	Vessel requirements	Prioritised fisheries
Demersal factory trawler	Length >39m Permit and quota for fishing with trawl for cod, haddock, saithe, north of 62°N Permit and quota for fishing with trawl for saithe, south of 62°N One or more vessels with licence for shrimp-trawl north of 62°N One vessel >53 m and equipped for fillet production One or more vessels equipped also for pelagic trawl	Cod, haddock, saithe with demersal trawl north of 62°N outside 12 nautical miles Saithe with demersal trawl south of 62°N outside 12 nautical miles Beaked redfish with pelagic/demersal trawl Greenland halibut with demersal trawl Shrimp trawl in the Barents Sea outside 12 nautical miles
Gillnet vessel fishing mainly in the North Sea	Length 28–40m Permit and quota for fishing with conventional gear (gillnet, longline, demersal seine) for cod south of 62°N Primary fishery for the vessel must be gillnet targeting cod in the North Sea	Cod, haddock, saithe with gillnet south of 62°N outside 12 nautical miles
Gillnet vessel fishing mainly in the Barents Sea	Length 28–40m Permit and quota for fishing with conventional gear (gillnet, longline, demersal seine) for cod, haddock, saithe north of 62°N Primary fishery for the vessel must be gillnet targeting cod in the Barents Sea and saithe on the fishing banks north of 62°N	Cod, haddock, saithe with gillnet north of 62°N outside 12 nautical miles Greenland halibut with gillnet
Longline and combined longline/gillnet vessel	Length >35m Permit and quota for fishing with conventional gear (gillnet, longline, demersal seine) for cod, haddock, saithe north of 62°N Two vessels with permit and quota for fishing with conventional gear (gillnet, longline, demersal seine) for saithe south of 62°N One vessel with permit and quota for fishing with conventional gear (gillnet, longline, demersal seine) for cod south of 62°N Primary fishery for the vessel must be longline targeting cod, haddock, ling, tusk, Greenland halibut and wolffish Two vessels fishing directly saithe with gillnet both north and south of 62°N One or more vessels with activity annually west of 4°W One or more vessels fishing cod, ling and tusk in the North Sea	Cod, haddock, saithe with longline/gillnet north of 62°N outside 12 nautical miles Cod, saithe with longline/gillnet south of 62°N outside 12 nautical miles Ling and tusk with longline north and south of 62°N outside 12 nautical miles Wolffish in the Barents Sea Greenland halibut with longline/gillnet Ling, tusk with longline/gillnet west of 4°W
Demersal seine/ purse seine vessel	Length >28m Permit and quota for fishing with conventional gear (gillnet, longline, demersal seine) for cod north of 62°N Permit and quota for fishing with purse seine for saithe north of 62°N Primary fisheries for the vessel must be with demersal seine for cod and with purse seine for saithe north of 62°N	Cod, haddock, with demersal seine north of 62°N outside 4 nautical miles Saithe with purse seine north of 62°N Norwegian Spring spawning herring with purse seine North Sea herring with purse seine Mackerel with purse seine

Category	Vessel requirements	Prioritised fisheries
Industry trawler (vessel targeting species primarily used for fish- meal production)	Licence for pelagic trawl Primary fisheries for the vessel must be with trawl for sandeel, Norwegian pout and blue whiting in the North Sea One vessel with permit and quota for fishing silver smelt with pelagic trawl north of 62°N	Sandeel with trawl in the North Sea/ south of 62°N Norwegian pout/blue whiting mixed fishery with trawl in the North Sea/ south of 62°N Saithe as retained bycatch in the North Sea/ south of 62°N trawl fishery Blue whiting with pelagic trawl outside 12 nautical miles Mackerel with pelagic trawl outside 12 nautical miles Norwegian Spring spawning herring with pelagic trawl outside 12 nautical miles North Sea herring with pelagic trawl outside 12 nautical miles North Sea sprat with pelagic trawl outside 12 nautical miles Capelin with pelagic trawl Silver smelt with pelagic trawl north of 62°N

Table A2. Vessel categories in the Coastal Reference Fleet. See Figure A1 for map of statistical areas

Category	Vessel requirements	Prioritised fisheries
Gillnet/longline vessels north Home harbours in statistical areas 03, 04, 05, 00, 06 & 07	Length 9–16m Home adresse and carries out most of its fishing in one of the areas described under the vessel category Active in the predominant coastal fisheries for the area Main fishing gear is gillnet/longline	Cod, haddock, saithe with gillnet/longline coastal north of 62°N Ling and tusk with gillnet/longline coastal north of 62°N Anglerfish with gillnet north of 62°N Greenland halibut coastal fishery with gillnet/longline north of 62°N
Gillnet/longline vessel south Home harbours in statistical areas 28, 08 & 09	Length 9–16m Home adresse and carries out most of its fishing in one of the areas described under the vessel category Active in the predominant coastal fisheries for the area Main fishing gear is gillnet/longline	Cod, haddock, saithe with gillnet/longline coastal south of 62°N Anglerfish with gillnet south of 62°N Mackerel coastal fishery with gillnet/jigging/other gears Wrasse pot fishery
Demersal seine vessel north Home harbour in statistical area 03	Length 9–16m Home adresse and carries out most of its fishing in one of the areas described under the vessel category Active in the predominant coastal fisheries for the area Main fishing gear is demersal seine	Cod, haddock, saithe with demersal seine coastal north of 62°N
Demersal seine vessel south Home harbour in statistical area 08	Length 9–16m Home adresse and carries out most of its fishing in one of the areas described under the vessel category Active in the predominant coastal fisheries for the area Main fishing gear is demersal seine	Cod, haddock, saithe with demersal seine coastal south of 62°N Mackerel coastal fishery with seine/other gears
Shrimp trawler – Skagerrak and North Sea Home harbours in statistical areas 08 & 09	Length 9–15m One vessel with length 15–28m Home adresse and carries out most of its fishing in one of the areas described under the vessel category Active in the coastal shrimp fishery Main fishing gear is shrimp trawl	Shrimp fishery in the Skagerrak and North Sea

Table A3. List of vessels in the High-Seas Reference Fleet between 2015 and 2019

Vessel category	2015	2016	2017	2018	2019
Demersal factory trawler	Andenesfisk 1 (LJWI) Havbryn (LDBT) Hermes (LLOP) Ramoen (LMLT) Vesttind (LLDH)	Andenesfisk 1 (LJWI) Havbryn (LDBT) Hermes (LLOP) Vesttind (LLDH)	Andenesfisk 1 (LJWI) Havbryn (LDBT) Hermes (LLOP) Ramoen (LDNV)	Havbryn (LDBT) Hermes (LLOP) Ramoen (LDNV)	Gadus Neptun (LDDG) Havbryn (LDBT) Hermes (LLOP) Ramoen (LDNV)
Gillnet vessel fishing mainly in the North Sea	Nesejenta (3WYO) Skjongholm (LHSQ)	Nesejenta (3WYO) Skjongholm (LHSQ)	Nesejenta (3WYO) Skjongholm (LHSQ)	Nesejenta (3WYO) Skjongholm (LHSQ)	Nesejenta (3WYO) Skjongholm (LHSQ)
Gillnet vessel fishing mainly in the Barents Sea	Kato (LLJC)	Kato (LLJC)	Kato (LLJC)	Kato (LLJC)	Kato (LLJC)
Longline/gillnet vessel	Carisma Viking (LLPZ) Nesbakk (LJZJ) O.Husby (LJQG) Vonar (LMCJ)	Carisma Viking (LLPZ) Nesbakk (LJZJ) O.Husby (LJQG) Vonar (LMCJ)	Atlantic (LIYX) Nesbakk (LJZJ) O.Husby (LJQG) Vonar (LMCJ)	Atlantic (LIYX) Nesbakk (LJZJ) O.Husby (LJQG) Vonar (LMCJ)	Atlantic (LIYX) Nesbakk (LJZJ) O.Husby (LJQG) Vonar (LMCJ)
Demersal /purse seine vessel	Hovden Viking (JWLM) Skagøysund (LMUR)	Hovden Viking (JWLM) Skagøysund (LMUR)	Kamilla Grande (JWLM) Skagøysund (LMUR)	Kamilla Grande (JWLM) Skagøysund (LMUR)	Hovden Viking (LEYN) Skagøysund (LMUR)
Industry trawler	Cetus (LLYM) Herøyfjord (LMHM)	Cetus (JXML)	Cetus (JXML) Håflu (LEQI)	Håflu (LEQI) Vikingbank (LLAS)	Cetus (LFFK) Håflu (LEQI) Vikingbank (LLAS)

Table A4. List of vessels in the Coastal Reference Fleet between 2015 and 2019. See Figure A1 for map of statistical areas

Category	**Statistical area**	**2015**	**2016**	**2017**	**2018**	**2019**
Gillnet/longline vessels north.	03	Solgløtt (LM2890)	Solgløtt (LM2890)	Solgløtt (LM2890)	Solgløtt (LM2890)	Solgløtt (LM2890)
	04	Odd Yngve (LM2864) Øyværing (LM8662)	Odd Yngve (LM2864) Øyværing (LM8662)	Odd Yngve (LM2864) Øyværing (LM8662)	Odd Yngve (LM2864) Øyværing (LK3925)	MT Senior (LG7408) Øyværing (LK3925)
	05	Ægir (LK5045) Vornesværing (LK5647)	Ægir (LK5045) Vornesværing (LK5647)	Ægir (LK5045) Vornesværing (LK5647)	Ægir (LK5045)	Ægir (LK5045) Braken (LM7459)
	00/05	T.Sivertsen (LK5948) Hellskjær (LM8308)	T.Sivertsen (LK5948) Hellskjær (LM8308)	T.Sivertsen (LK5376) Hellskjær (LM8308)	T.Sivertsen (LK5376) Hellskjær (LM8308)	T.Sivertsen (LK5376)
	00	Rånes Viking (LK5016) Økssund (LK6737)	Rånes Viking (LK5016) Økssund (LK6737)	Rånes Viking (LK5016) Økssund (LK6737)	Rånes Viking (LK5016) Økssund (LK6737)	Rånes Viking (LK5016) Økssund (LK6737)
	06	Haldorson (LK4789)	Haldorson (LK4789)	Haldorson (LK4789)	Haldorson (LK4789)	Haldorson (LK4789)
	07	Tramsegg (LK7141) Haaværbuen (LM5498) Øygutt (LK5160)	Tramsegg (LK7141) Haaværbuen (LM5498) Leon Olai (LK2759)	Tramsegg (LK7141) Sørhav (LG4010)	Tramsegg (LG3690) Sørhav (LG4010)	Tramsegg (LG3690) Sørhav (LG4010)
Demersal seine vessel north	03	Charmi (LK3293)	Charmi (LK3293)		Kristian Gerhard (LK7556)	Kristian Gerhard (LK7556)
Gillnet/longline vessel south	28	Vester Junior LM5970) Britt Evelyn (LK6966)	Vester Junior LM5970) Britt Evelyn (LK6966)	Vester Junior LM5970) Britt Evelyn (LK6966)	Vester Junior LM5970) Britt Evelyn (LK6966)	Vester Junior LM5970) Britt Evelyn (LK6966)
	08	Austbris (LK9305) Ramona (LK6606) Repsøy (LK3270)	Austbris (LK9305) Ramona (LK6606) Repsøy (LK3270)	Austbris (LK9305) Ramona (LK6606) Vicma (LG9311)	Austbris (LK9305) Ramona (LK6606) Eggøy (LM8940)	Trellevik (LG4914) Fjorden (LK6326) Eggøy (LM8940)
	09	Skogsøyjenta (LK5485) Vesleper (LM7915)	Skogsøyjenta (LK5485)	Skogsøyjenta (LK5485)	Skogsøyjenta (LK5485)	Skogsøyjenta (LK5485)
Demersal seine vessel south	08	Molinergutt (LG7405)	Molinergutt (LG7405)	Molinergutt (LG7405)	Molinergutt (LG7405)	Molinergutt (LG7405)
Shrimp trawler (9-15m)	09	Brattholm (LK7238) Tormo (LM3995)	Brattholm (LK7238) Tormo (LM3995) Mostein (LK5352)	Brattholm (LK7238) Tormo (LM3995) Mostein (LK5352)	Grepan Junior (LK5485) Tormo (LM3995) Mostein (LK5352)	Brattholm (LH2820) Tormo (LM3995)
Shrimp trawler (15-28m)	08/09					Guldringnes (LKZZ)

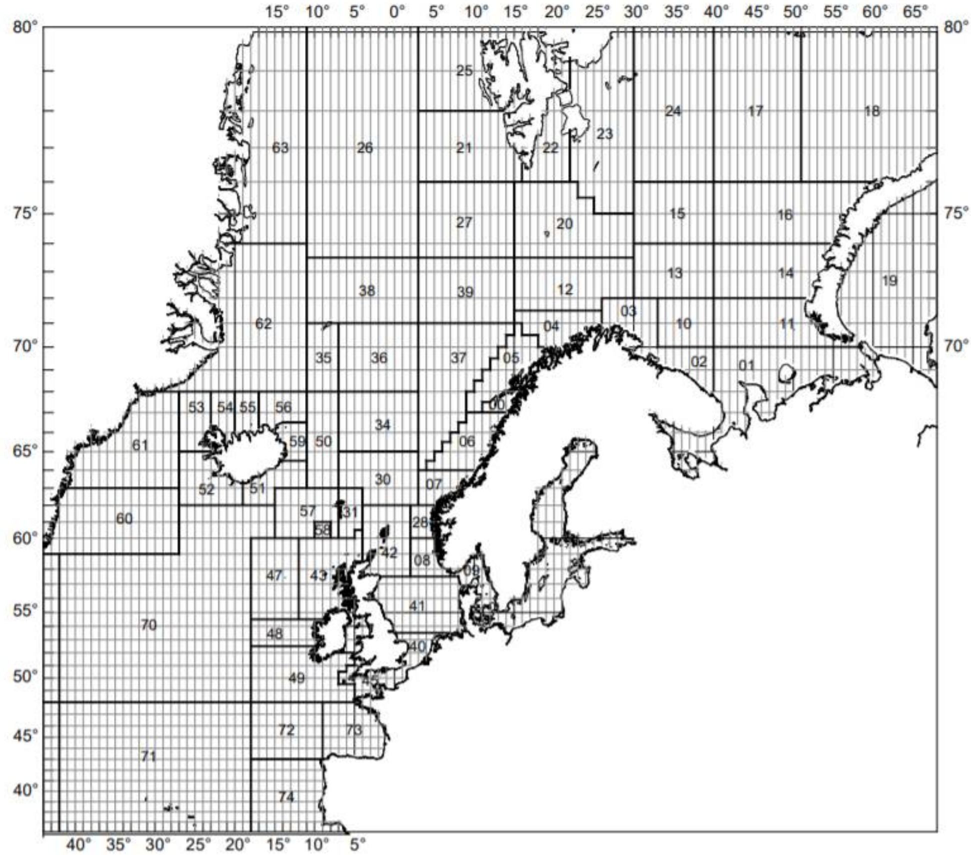


Figure A1. Map of statistical areas defined by the Norwegian Directorate of Fisheries

5.2 – Appendix B: Sampling protocols

Table B1. Protocol for catch registration and sampling in the High-Seas Reference Fleet

Gear type	Catch registration	Sampling
Demersal trawl	Every haul – the processed (landed) catch is registered and bycatch of seabirds, sea-mammals and seldom fish species (e.g. porbeagle and basking shark). From 2019 registering bycatch of corals and sponges is also included in the procedure. One haul every other day - total catch is registered, including all bycatch species and discards of both commercial and bycatch species. From 2019 discards are registered separately from the retained catch per species that is processed for fishmeal.	One haul every other day – length and weight measurements are taken of up to 20 individuals of all species in the catch, both landed and from discards One haul per week – Otolith samples are taken for important demersal species
Shrimp trawl	Every haul – the processed (landed) catch is registered and bycatch of seabirds, sea-mammals and seldom fish species (e.g. porbeagle and basking shark). From 2019 registering bycatch of corals and sponges is also included in the procedure. One haul every other day - total catch is estimated from 3 basket samples from the catch and registered, including all bycatch species and discards of both commercial and bycatch species. From 2019 discards are registered separately from the retained catch per species that is processed for fishmeal.	One haul every other day – length and weight measurements are taken of up to 50 individuals of all species in the catch, both landed and from discards
Demersal seine	Every other haul – the processed (landed) catch is registered and bycatch of seabirds, sea-mammals and seldom fish species (e.g. porbeagle and basking shark). One haul every other day - total catch is registered, including all bycatch species and discards of both commercial and bycatch species.	One haul every other day – length and weight measurements are taken of up to 20 individuals of all species in the catch, both landed and from discards One haul per week – Otolith samples are taken for important demersal species
Pelagic trawl and purse seine	Every other haul/cast – the processed (landed) catch is registered and bycatch of seabirds, sea-mammals and rare fish species (e.g. porbeagle and basking shark). Every alternate haul/cast - total catch is registered, including all bycatch species and discards of both commercial and bycatch species. End of trip – if the onboard pumping of the catch is a closed system. Total catch, including bycatch species. Hauls/casts with zero catch or slipping of all/part of the catch is also registered	Every other haul/cast –samples length and weight measurements for all species in the catch. Number of individuals in a sample dependent upon the species Every other haul/cast –frozen sample of target species for length/age determination and other important variables. For some pelagic species frozen samples are taken for each catch. One catch per week – Otolith samples are taken for important demersal species
Industrial trawl (Target species: Sandeel, Norwegian pout & Blue whiting)	Every haul – the landed catch is separated in to catch to consume and catch that is pumped into the holding tanks for fish-meal production, and registered by species. Bycatch of seabirds, sea-mammals and seldom fish species (e.g. porbeagle and basking shark) is also registered. From 20One9 registering bycatch of corals and sponges is also included in the procedure. One haul every other day - total catch is registered, including all bycatch species and discards of both commercial and bycatch species. Species composition catch that is pumped into the holding tanks is estimated from 3 basket samples of following the IMR sampling procedure for catch sampling.	Every other haul – frozen sample of some target species for length/age determination and other important variables. For some species frozen samples are taken for each catch. One haul every other day – length and weight measurements are taken of samples of all species in the catch, both landed and from discards. The number of individuals in a sample dependent upon species. One haul per week – Otolith samples are taken for important demersal species

Gear type	Catch registration	Sampling
Long-line/gillnet	Every daily catch – the processed (landed) catch is registered and bycatch of seabirds, sea-mammals and seldom fish species (e.g. porbeagle and basking shark). Effort is recorded in number of hooks/gillnets, but not soak time. Every other day – for a representative portion of the total gear hauled that day (approximately 16,000 hooks or 100 gillnets), total catch is registered, including all bycatch species and discards of both commercial and bycatch species. Effort is recorded in number of hooks/gillnets and soak time.	One haul every other day – length and weight measurements are taken of up to 20 individuals of all species in the catch, both landed and from discards One haul per week – Otolith samples are taken for important demersal species

Table B2. Protocol for catch registration and sampling in the Coastal Reference Fleet

Gear type	Catch registration	Sampling
All gear types	Each day – total catch is registered, including all bycatch species and discards of both commercial and bycatch species. Shrimp trawl – from 2019 registering bycatch of corals and sponges is also included in the procedure. Splitting the catch – if the day's catch is taken from multiple fishing operations from different depths, fishing area or different gear types, then the catch should be split and registered separately. For example, two gillnets used the same day with different mesh-sizes and set at different depths.	One catch per week– length and weight measurements are taken of up to 20 individuals for each species in the catch, both landed and from discards. Otolith samples are taken for important demersal species

HAVFORSKNINGSINSTITUTTET

Postboks 1870 Nordnes
5817 Bergen
E-post: post@hi.no
www.hi.no

REPORT II

FISKERIDIREKTORATET

Rapport

Estimating the size distribution of reported catches on-board factory vessels – Issues with using data from the production process

Thomas L Clegg, Geir Blom, Kotaro Ono, Kjell Nedreaas, oktober 2021

Livet i havet – vårt felles ansvar

Rapport

Estimating the size distribution of reported catches on-board factory vessels – Issues with using data from the production process

Årstall

2021

Ansvarlig avdeling:

Ressursavdelingen og Statistikkavdelingen (Fiskeridirektoratet)/ Fiskeridynamikk (Havforskningsinstituttet)

Emneord:

High-grading, bycatch, discard, fishmeal, factory vessels, length-distribution, bias, precision

Arkivsaksnummer:

Dato utgitt:

01.11.2021

Totalt antall sider:

22

Saksansvarlig:

Thomas L. Clegg, Geir Blom, Kotaro Ono og Kjell Nedreaas

ISBN: ISBN-13: 9788292075098

Sammendrag

‘High-grading’ er praksisen med å kaste ut fangster med lavere verdi for å få plass til fangster med høyere verdi. Det er nødvendig å forstå omfanget av urapportert utkast og dets variasjon for å forbedre bestandsvurderinger og forvaltningsbeslutninger. Der det ikke er direkte observasjoner av utkast av fisk, finnes det metoder for å estimere ‘high-grading’ ved å sammenligne størrelsesfordelinger av totalfangst (før sortering) og landet fangst (etter sortering), men per i dag er det ikke tilgjengelig en tilstrekkelig detaljert datakilde for den landete delen. Denne rapporten presenterer data fra to pilotstudier for å undersøke egnetheten til data samlet inn i forbindelse med om bord produksjon for å beskrive størrelsesfordelingen av rapporterte fangster. Produksjonsrapportene vi mottok inneholdt aggregerte data, der individuelle vekter av en fiskeart var aggregert i grove, overlappende vektintervall. Slike data resulterer i et stort tap av informasjon. Bruk av statistiske prosedyrer for å få et mer detaljert bilde av størrelsesfordelingene av fisk, vil introdusere enda mer usikkerhet i et allerede usikkert datamateriale som potensielt kan føre til ikke-signfikante resultater og som kan introdusere ukjente skjevheter. Vi konkluderer med at det er nødvendig å bruke rådataene bak aggregerte produksjonsrapporter. Rådataene består av registrert produktvekt av ulike fiskearter på individnivå. I denne forbindelsen diskuterer vi de logistiske og statistiske problemene som produksjonsdata inkludert observasjoner på individnivå kan introdusere.

Summary

High-grading is the practice of discarding lower value catches to make space for catches with higher value. It is necessary to understand the extent and variation in these unreported discards to improve stock assessments and management decisions. Where discards are not directly observed, a proposed methodology for estimating high-grading involves comparing size distributions of total catches (before sorting) and landed catches (after sorting), but we have yet to identify a suitable data source for the landed portion. This report presents data from two pilot studies exploring the suitability of data gathered during the on-board factory production process for describing the size distribution of reported catches. We received these data in a summarised report, where individual fish weights are aggregated into coarse, overlapping size grades. This summarised form results in a large loss of information. Applying the necessary statistical procedures to get a more detailed picture of fish size distributions would introduce even more uncertainty into an already uncertain estimation, potentially leading to non-signficant results, and can introduce unknown biases. We conclude that it is necessary to use the raw data behind summarised reports which provide data on individual fish. To this end, we address the logistical and statistical issues posed by production data including individual fish observations.

Rapport

Estimating the size distribution of reported catches on-board factory vessels

Innholdsfortegnelse

Tables

Figures

1. Introduction

High-grading is the practice of discarding lower value catches to make space for catches with higher value (Batsleer *et al.*, 2015). High-grading of commercial species is typically size-based, influenced by the minimum landing size or market prices favouring larger individuals. Discarding is illegal under a landing obligation and therefore results in misleading catch statistics. For example, neglecting discards of small fish in stock assessments can mask strong incoming year-classes (Punt *et al.*, 2006). A good knowledge of high-grading is therefore necessary to improve stock assessments (Dickey-Collas *et al.*, 2007; Perretti *et al.*, 2020) and have a more realistic understanding of the environmental impact of the fishery.

The Norwegian discard ban was implemented in 1987 to mitigate against the emerging practice of high-grading (Gullestad *et al.*, 2015). Since then, a suite of accompanying measures has been developed, known collectively as the 'discard ban package' (Gullestad *et al.*, 2015), to build a more comprehensive policy to reduce discards. Fishers are incentivised to land illegal catches through compensation; avoidance is actively encouraged through legal obligations to move away from high-risk areas; and fishing gears are being constantly developed to improve selectivity. However, it is known that discarding still occurs through direct observations by the Norwegian Coast Guard that have resulted in prosecutions. There are also various studies estimating historical high-grading in the Barents Sea trawl fishery for cod (*Gadus morhua*). In the absence of direct observations, empirical gear selectivity curves have been applied to fisheries-independent sampling to simulate commercial fishing (McBride and Fotland, 1996; Dingsør, 2001). Breivik *et al.* (2017) estimated historical bycatches of cod in the Barents Sea shrimp fishery using the Directorate of Fisheries Monitoring and Surveillance Service (MSS), a programme which hires or joins fishing vessels. However, this sampling is focussed on species of concern (e.g. juvenile cod bycatch) and does not regularly record size measurements of species.

More recently, the Norwegian Reference Fleet have provided direct sampling of species (Clegg and Williams, 2020), including size measurements which allows for more direct methods for quantifying high-grading. In the coastal segment of the Norwegian Reference Fleet, vessels sample discarded and landed portions of the catch separately, which enables high-grading to be readily identified (Berg and Nedreaas, 2020). However, in offshore fisheries, the Norwegian Reference Fleet have only sampled discards and landed catches separately since 2019. Prior to this, vessels only sampled total catches (i.e. discards and landed catches combined). In this situation, we can still estimate fishery-level high-grading by comparing the size distributions of observed total catches with those of landed catches (e.g. Pálsson, 2003). However, we have yet to identify a reliable source of size-based data on landed catches to make such a comparison.

There are numerous size-based data sources in the mandatory catch reporting framework, but unlike the Norwegian Reference Fleet, official catch reports do not offer an adequate data resolution for quantifying high-grading. Daily logbooks have a high spatial and temporal resolution, but do not record size-based information on catches. Sales notes are generated once a vessel lands the catches after each trip. They do include size-based information, but only as a summary of an entire trip and in course market-defined size grades. In one trip, a vessel spans many statistical areas over a period of weeks or even months, meaning trip-level resolution is insufficient to understand spatial and temporal patterns in high-grading.

For offshore fisheries prior to 2019, where discards were not reported explicitly, the other scientific sampling programmes did not provide information on unreported portions of the catch. Coast Guard sampling is done by enforcement officers, meaning that fishers will not discard in their presence. A port intercept sampling programme only samples coastal vessels landing fresh fish north of 62°N latitude (Hirst *et al.*, 2004), whilst a newer mandatory self-sampling programme covers offshore fisheries for selected pelagic species (Stenevik *et al.*, 2020).

This report focuses on estimating high-grading in offshore Norwegian fisheries prior to 2019, where the Norwegian Reference Fleet did not sample landed and discarded catches separately. We present two pilot studies which trialled a potential source of size-based data of reported catches generated by on-board factory production systems. Vessels constantly monitor productivity in the on-board factory to inform fishing strategy and to keep a record of catches on board, which also contributes to mandatory reporting. At present, these data are the most detailed source of size-based information for retained catches. We requested summarised reports (hereafter referred to as production reports) from vessels in two fisheries to explore their viability for use in quantifying high-grading. We describe the production process on-board factory vessels and how production reports are generated. We then explore the statistical properties of the data to identify issues that restrict their viability for quantifying high-grading and discuss how these issues can be rectified for future studies.

2. Description of the production process on board factory vessels

As fish enter the on-board factory immediately after hauling, crew sort through the catches to decide which fish to process (Figure 1). Fish can either be processed into a range of products, discarded, or converted into highly-processed products such as fishmeal or ensilage.

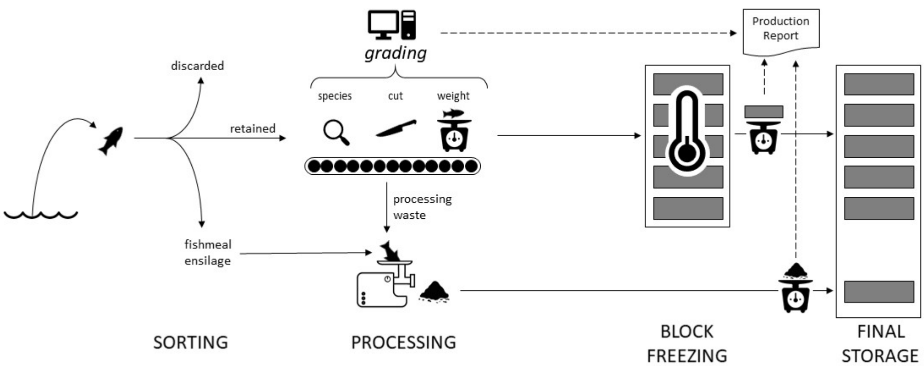


Figure 1: Generalised description of the production process on-board a factory vessel. The specific process can vary depending on factory setup.

During the production process, individual fish are cut for the desired product and then weighed. This means there is no information on the round weight (original live weight) of fishes. The species is either registered manually, inferred from the route taken through the factory, or more recently by image recognition software. Using automatic conveyor belt systems, all this information is combined to grade each fish based on the species, product and processed weight. Graded products are held in temporary storage tanks, where they are then frozen in standardised blocks and transferred to the final storage hold.

The production process generates no information about unreported catches, which includes illegally discarded fish and fish processed as fishmeal (NOU, 2019). We therefore cannot directly quantify the scale of high-grading due to illegal discarding, and how much is legally landed as fishmeal. However, if a vessel has on-board fishmeal production, then we can assume that all unwanted fish are processed into fishmeal, meaning unreported catches on those vessels are not a result of high-grading. The majority of vessels do not have fishmeal production facilities on-board, but new vessels are increasingly installing them. If the vessel cannot produce fishmeal, unwanted catches can be frozen whole as mixed species to be delivered to production facilities on land.

To monitor production and assist in mandatory reporting, summary reports can be generated for any given time period. The report aggregates the number and weight (measured or estimated) of frozen blocks from the final output stage to provide the total weight of each product and grade. The report can also be supplemented with additional information from the grading process, such as the average individual weight in each grade, which is used for estimating the total number of pieces of each product.

If the factory exceeds production capacity, crew can bypass the automated steps in the process to speed up production. This can result in certain information not being recorded. For example, we observed one system where crew record species and product by dropping incoming fish into defined hoppers. The hoppers then drop fish onto a conveyor belt at defined intervals, which passes over a weighing scale to record its weight. If a catch is large and dominated by a certain species and grade, the defined hopper may not drop fish fast enough, causing a backlog in the system. Crew can avoid this by storing those fish in baskets, then manually adding them to the correct temporary storage tanks after production has calmed, knowing the weight will be registered after freezing. This solution creates a risk that not all fish are recorded in the grading machine. Furthermore, these manually graded fish will likely differ in size to those automatically graded.

Using production reports as a source of information on high-grading assumes that all fish entering the processing stage are ultimately landed and reported. Vessels are legally required to maintain an accurate record of all catches stored on the vessel. These catch diaries are filled out based upon production data described above. The catch diaries can be inspected at any time at sea by enforcement authorities. They should then match with the sales of catches after returning to land. Substantial inconsistencies between the catch diary and landing report will be automatically flagged and can be investigated. Norway has previously been ranked highest in fisheries compliance globally (Pitcher *et al.*, 2009), owing to efficient regulations and enforcement, and a broad willingness for fishers to comply (Gezelius, 2006). Therefore, for the purposes of a study on high-grading, we can assume that production reports are representative of reported catches with respect to reliability of reporting.

3. Pilot studies – production report data requests

We ran two pilot studies to investigate the utility of production report data for estimating size distributions of species in reported catches in offshore fisheries. For both pilot studies, we contacted a selection of vessels to voluntarily provide daily production reports. The letter stated that data will be used for estimating unreported bycatches in the fishery, and that raw data will be treated as sensitive in accordance to data privacy laws. The letter also stated that vessels will not be prosecuted on the basis of submitted data, and that published materials will be aggregated and anonymised such that individual vessels cannot be identified.

The first pilot study was the Barents Sea trawl and autoline fisheries (Figure 2) for cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) in 2012. We randomly selected 10 trawl vessels and five autoline vessels for each annual quarter, with a probability weighted by total reported catches in the previous year. The selection for each annual quarter was independent, such that it was possible to select a vessel for multiple quarters. We contacted a total of 45 vessels, of which 18 vessels cooperated by providing the requested data. In four cases, reports were summarised over periods longer than a day, ranging from 14-25 days.

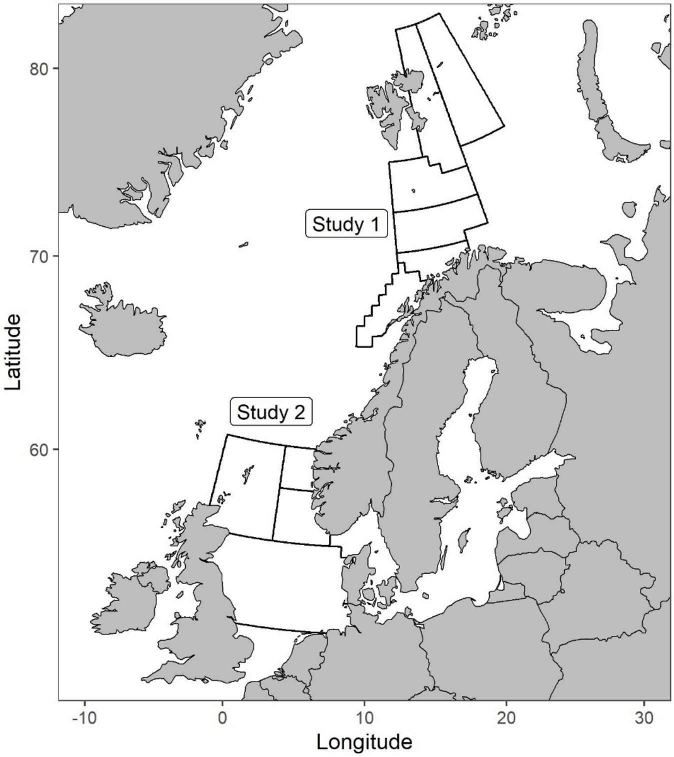


Figure 2: Map of study fisheries showing statistical areas included in the two pilot studies.

The second pilot study was the North Sea trawl fishery for saithe (*Pollachius virens*) (Figure 2) in 2018. We replicated the request from the first pilot study, with 13 vessels selected for each annual quarter, again with the possibility of multiple requests across quarters. We additionally requested all vessels to provide supplementary information from the grading machine for each daily production report, namely the mean weight of individual fish in each size grade. Of the 31 vessels contacted, 14 responded with production report data. All vessels returned reports for individual days. Only three vessels provided the supplementary information on mean weight from the grading machine data, which were summarised over periods ranging from one to 32 days.

We additionally requested four vessels to provide the raw data from the grading machine that is used to generate the supplementary information on mean weight. However, no vessels fulfilled this request.

3.1. Issues with data requests

In both pilot studies, data were typically provided as printable reports structured to improve readability, which hindered data entry. Furthermore, many vessels provided reports in PDF or paper formats, requiring an additional step to digitise and extract the values.

A large number of companies did not respond to our request. Unless a company actively objected to the request, we cannot determine if the rejection was intentional objection or due to neglect. This is important to understand due to the implications on sampling biases. If companies object to the request for similar reasons (e.g. high risk of prosecution), and furthermore, those characteristics differ from the general population, then the final sample may not be representative of the fishery. The request required a large administrative task to access, filter, and compile the dataset, then send it. For this reason, it is possible that many of the non-responses may be neglectful and could therefore be deemed random and not introduce bias.

4. Estimating unreported catches using production reports

To identify high-grading in a fishery where we do not have direct observations of discarded catches, we can infer it by comparing the size distribution of total catches with that of landed catches. If the two size distributions are different, then we attribute this to high-grading (Pálsson, 2003). It is important to highlight that we cannot determine if these unreported catches were discarded or converted to fishmeal.

Size distributions of total catches are available from the Norwegian Reference Fleet catch sampling programme. Vessels participating in the programme regularly record the length and weight of fish from samples of total catches (see Clegg and Williams (2020) for detailed sampling protocols), which can be aggregated across vessels to give a high-resolution size distribution of species for any given temporal or spatial scale.

In their standard format, production report data are not comparable to the size distributions generated from individual fish measurements by the Norwegian Reference Fleet. Production reports summarise total weight of products for each species in different size grades (Table 1). Therefore, a comparison would require the transformation of one or both datasets to standardise them.

One solution would be to reduce the resolution of Norwegian Reference Fleet data by aggregating individual fishes into size grades that match production reports. Even ignoring the complications in this process, reducing data resolution is highly undesirable in principle. It would be wasteful for the effort and money spent in sampling individual fishes and would only increase uncertainty.

Alternatively, we could infer the underlying size distribution of species from aggregated production reports. This would offer an estimated number of fish in smaller size intervals comparable with Norwegian Reference Fleet sampling. In this section, we present potential methods for inferring underlying size distributions, and explain why the data structure of production reports creates issues with this approach.

4.1. Round weight vs processed weight

The first obstacle met is the difference in measures of weight between processed catches in production reports and total catches recorded by the Norwegian Reference Fleet. Where fish are processed on-board the vessel, then fish weights are recorded after processing and

Table 1: Extract from a daily production report listing total production in a 24-hour period in January 2012 in statistical area 12.

Species	Product	Grade (kg)	Total product weight (kg)
Cod *Gadus morhua*	Fillet with bone with skin	0.45–0.91	1222
		mix	1364
	Headed & gutted	1–2.5	1749
		2.5–5	2325
		>5	144
Golden redfish *Sebastes norvegicus*	Headed & gutted	>1	79
	Headed & gutted (Japan cut)	≤1	47
Haddock *Melanogrammus aeglefinus*	Fillet without bone with skin	mix	819
	Headed & gutted	≤0.8	1260
		>0.8	1127
Halibut *Hippoglossus hippoglossus*	Headed & gutted	≤6	15
Spotted wolffish *Anarchichas minor*	Headed & gutted	≤3	43
Tusk *Brosme brosme*	Headed & gutted	≤1	36

converted back into round weight (the weight of the fish when it is taken from the water) using official conversion factors calculated by the Norwegian Directorate of Fisheries (Norwegian Directorate of Fisheries, 2021). The conversion factors are based on sampling on-board active fishing vessels, where fish are weighed before and after processing to estimate the average weight lost from production for each product. Conversion factors are published as annual mean values for all areas and are published without estimation error and are updated intermittently.

4.2. Weight-based vs. length-based assessment

Size-based stock assessments typically structure fish populations by age or length. The Norwegian Reference Fleet gather data on weight, length and age of individual fishes, providing usable data for stock assessments. Comparatively, production reports only include fish weight and would therefore require a conversion to either length or age to be comparable to data on total catches and subsequently useful for stock assessments.

4.3. Parametric approach

A parametric approach to inferring an underlying size distribution involves an assumption that observations come from a known distribution, which is described by a fixed set of parameters. Simpler distribution fitting methods in R (e.g. *MASS::fitdistr*; Venables and Ripley, 2002) assume that observations are known without error. This is not true for observations from production reports, which only offer the total weight of fish in a defined size range (Table 1). Fortunately, more advanced methods of distribution fitting can account for such uncertainties, such as the *fitdistrplus* package in R (Delignette-Muller and Dutang, 2015). This expands the distribution fitting functions to accommodate both interval and censored observations, which are key characteristics of production reports (Table 1).

4.4. Interval and censored size grades

In production reports, fish size grades are reported in intervals, such that we only know that fish were within a defined weight range. The largest and smallest size grades are typically censored, meaning that the size range is only partially known. Left-censored size grades (i.e. fish below a certain size) are limited to positive values, which is reflected in the appropriate distributions for weight data, namely gamma and log-normal distributions. However, these distributions can include values approaching zero. It is extremely unlikely to observe fish with a size approaching zero in catches due to size selective fishing gears and avoidance strategies (Reid *et al.*, 2019).

Right-censored size grades (i.e. fish above a certain size) include the largest fish caught in the fishery. The most likely reasons for unreported large fishes are either discarding of damaged individuals, or illegal sale. A right-hand censored size grade has no theoretical limit so if left undefined, the distribution fitting functions will estimate the limit based on the estimated parameter for the distribution. This limit could also be defined empirically, based on the largest fish observed in the Norwegian Reference Fleet if it improved model fit quality or convergence.

The *fitdistrplus* package in R (Delignette-Muller and Dutang, 2015) can account for censored size grades by estimating the cumulative distribution function for censored observations instead of probability density function for non-censored observations. Assessing the quality of model fit is not a simple comparison, but instead requires a judgement based on a suite of statistical tests and the visual inspection of graphical outputs. This is further complicated by the reason for fitting a distribution. In our situation, we want to identify if high-grading is occurring, which we expect to be size-based, such that small fish are more likely to be misreported. Any differences in size distributions between total and reported catches are therefore more likely in the left tail of the distribution. Fortunately, assessing the quality of distribution fitting can focus on this portion of the distribution at the expense of the right tail (larger fish) (Delignette-Muller and Dutang, 2015).

In extreme cases, a portion of catches may not be graded at all. This can be a result of market demands, grading errors, or damaged products. We have no information on whether this grade allocation is biased with respect to size, but an assumption of no bias will allow mixed grades to be removed from an analysis.

4.5. Estimating the number of fish in each size grade

To fit a distribution to fish size observations using censored data, we must first know the number of fish in each size grade. If only the total catch weight is reported for each size grade, we must estimate the number of fish by dividing the total weight by the average individual weight in each grade. An estimation of numbers of fish in each size grade based upon empirical data is desirable over using the midpoint of each size grade.

In the absence of an empirical estimate of the average individual weight, we must assume it to be the midpoint of each size grade. This assumes that size grades are independent of each other, and that in each size grade, observed weights are normally distributed and centred around the midpoint. In reality, size grades are not independent; they are an arbitrary division of a larger size distribution. Size grades in the left tail of the distribution will contain increasing numbers of larger fish, whilst those in the right tail will contain decreasing numbers, meaning that fish sizes in one size grade are seldom normally distributed. Considering in addition that larger fish contribute more weight, it is likely that the midpoint is a poor estimator of average individual fish weight in each size grade.

In the first pilot study, we did not have any empirical knowledge of average individual weight, so we must assume that it is the midpoint of each size grade. In the second study, some vessels provided the mean individual weight of fish in supplementary data from the grading machine. Using these additional data from the second study, we can demonstrate the importance of using empirical estimates of average individual weight. We first calculate the number of fish in each size grade using both the assumed average individual size (midpoint) and the empirical estimate (supplementary grading data). We can then calculate the percentage error introduced from assuming the average individual size:

$$Error_{j,k}\ (\%) = \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{x}_{ijk} - x_{ijk}}{x_{ijk}} \times 100 \qquad (1)$$

where for an individual production report, i for species, j and size grade, k, $\hat{x}_{ijk}$ is the assumed number of fish based upon the midpoint of the size grade and x_{ijk} is the estimated number of fish taken from supplementary data. For right-censored size grades, we assumed the upper limit was the largest fish observed by the Norwegian Reference Fleet.

Figure 3 shows that assuming the average individual weight of fish in each size grade is the midpoint of that grade results in large biases in estimates of total number of fish for all size grades. Furthermore, there is a strong trend across all three species to overestimate the number of fish in left-censored size grades when using the midpoint method. An overestimation of small fish in reported catches will mask the true scale of high-grading when comparing with observations of total catches by the Norwegian Reference Fleet. This issue is worsened by the underestimation of numbers of larger fishes. When fitting a parametric distribution to the size data, an over- and underestimation of small and large fish respectively will force the best-fitting distribution to be more positively skewed. Given that we expect discarding to be relatively low in Norwegian fisheries (Pérez Roda *et al.*, 2019; Gilman *et al.*, 2020), there is an increased risk that an underestimation of high-grading could conclude that it does not occur.

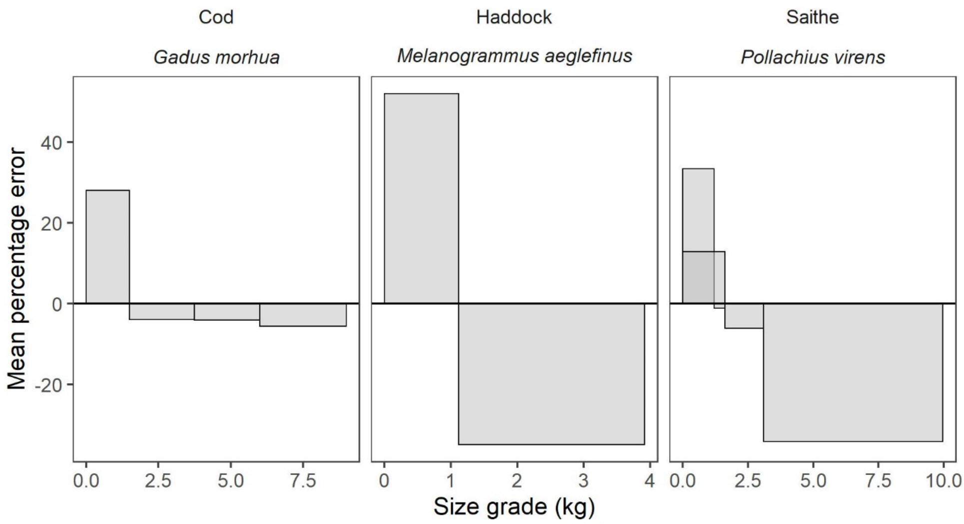


Figure 3. Error in estimated number of fish from assuming the average individual weight is the midpoint of each size grade. Each bar represents a size grade, which may overlap (e.g. Saithe). Left- and right-most grades are censored. Censored size grades were given assumed limits to allow for plotting (left-censored = zero; right-censored = largest individual observed by the Norwegian Reference Fleet.

4.6. Variations in grading intervals

There is no standardisation of grading intervals across the fishing fleet, as intervals vary depending on business strategy, market trends and catch composition. This makes it impossible to aggregate production reports on any level for many species without avoiding overlapping size grades. We could select only those reports with matching grading intervals to create a reduced dataset which could be aggregated. However, we do not know enough about the detailed reasons for grading decisions to understand whether removing certain grading intervals would bias sampling.

4.7. Coarse grading

For some species such as haddock, grading is limited to a small number of categories. For example, across both pilot studies 58.1% of daily production reports for haddock used only two size grades. Furthermore, both of these grades are censored (i.e. either larger or smaller than a defined weight). Similarly, 16.4% of daily production reports contained only one grade for either cod or haddock. In these situations, there is simply not enough information in the summarised reports to estimate the underlying distribution.

With wide, left-censored size grades, it is possible that only one size grade describes the left tail of the distribution. In this situation, it is likely that the distribution of fish sizes will be best described as an exponential decrease. Candidate distributions that typically describe fish sizes, such as the gamma and log-normal distributions, can approach or become exponential given certain parameterisations. An exponential distribution is not suitable for describing the size distribution of fish caught by trawl and autoline fishing gears, which are selective for larger individuals (Reid *et al.*, 2019).

4.8. Multimodal size distributions

We cannot determine if a multimodal distribution should be fitted based on the information from production reports (Table 1), due to grading being very coarse and censored. Fitting unimodal (single-peaked) distributions to infer the underlying fish size distribution ignores the possibility of multimodal size distributions (multiple peaks). Multimodal size distributions are common in fish populations due to a wide range of biotic and abiotic factors (Huston and Deangelis, 1987). These occur either as multiple factors influencing the same population, or the unintentional combining of distinct groups of fishes (e.g. populations or different life history stages). Fish are typically seasonal spawners, meaning they are born in a discrete time period and grow as a cohort. However, individual growth rates will vary due to a wide range of factors such as genetics, environment and food availability (Huston and Deangelis, 1987). Habitat associations at different life stages can separate fish size on a fine spatial scale and cause seasonal variations in size distributions (Methratta and Link, 2007). It is difficult to capture these factors for all species using a single stratification system.

Figure 4 shows two examples of multimodal size distributions in total catches sampled from both Norwegian Reference Fleet and Coast Guard inspections. In such examples, we could be observing multiple cohorts, sex-dependent growth differences or overlapping populations. Whilst we can speculate any number of underlying causes here, the main point is that fitting a unimodal distribution to these two examples would skew the peak towards the major modes. Furthermore, this issue is focused around the minimum landing size of the species, which is the focal area for quantifying high-grading.

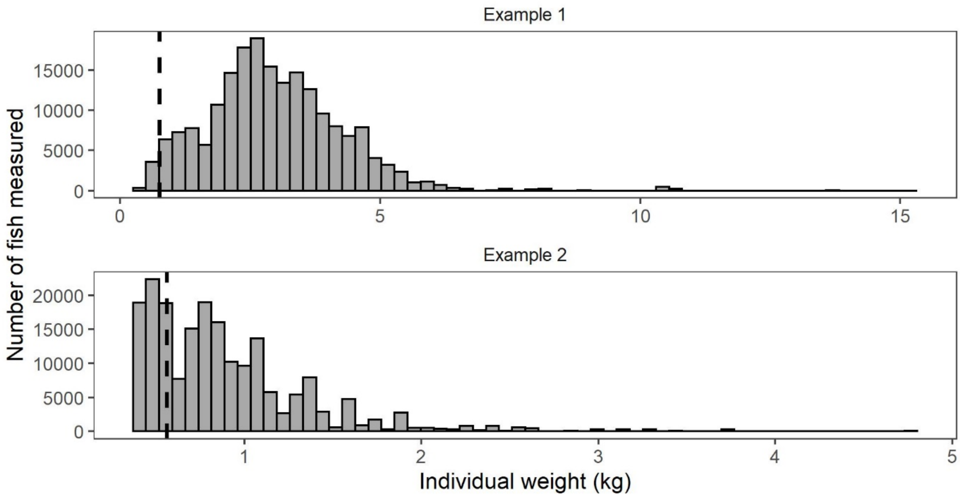


Figure 4. Examples of multimodal fish size distributions using data from the Norwegian Reference Fleet sampling and Coast Guard inspections. Dashed line indicates minimum landing size (converted from length). Example 1: Cod in the Barents Sea trawl fishery, statistical area 23, quarter 4; Example 2: Saithe in the North Sea trawl fishery, statistical area 08, quarter 3.

4.9. Non-parametric approach

Parametric approaches lack the ability to describe more complex distributions for which we do not know the underlying assumptions, the largest of these being multimodal distributions. On the other hand, non-parametric models make no assumptions about the distribution of data. As a result, non-parametric models need more data to understand the underlying functions, or more importantly, they need more information.

A non-parametric approach to estimating the underlying size distribution involves the ungrouping of coarsely aggregated data into smaller intervals (Rizzi *et al.*, 2015). A comparison of methods for ungrouping coarsely aggregated data by Rizzi *et al.* (2016) identified the penalized composite link model as the most efficient method for very coarsely aggregated data. The penalized composite link model can account for censored size grades but requires an estimation of the size of the right-censored grades (left-censored grades are limited at zero). In epidemiological applications, a right-censored age group is more easily limited to the oldest known age (Rizzi *et al.*, 2016), which is relatively easy to estimate. However, it is more difficult to estimate the largest fish in the fished population, as we only have information on those fish caught in samples rather than census information.

The largest obstacle to using a penalized composite link model for estimating the underlying size distribution is the need for sequential intervals in the raw count data. However, fish are graded depending on a wide range of factors including product, market demand and catch composition. Therefore, even on the level of a single haul or product, there is seldom a sequential grading system (Table 1). Even if a sequential grading system was used for each haul, product or time period, we would need to fit individual models to each haul, product or time period independently, reducing the data available for analysis and increasing risk of uncertainty and bias in estimates when combining the model outputs.

5. Discussion

This report has explored the statistical properties of data generated by production reports on board factory fishing vessels, which were obtained through two pilot studies to determine their utility for quantifying high-grading in offshore fisheries where scientific sampling of reported catches is unavailable.

5.1. Data collection

A large number of companies failed to provide production reports upon request, which could result in non-response biases. There were some cases of apprehension towards the intended use of the data, despite a clear definition of intentions and reassurance of safe data handling protocols. Some companies offered that independent on-board observers could collect the required data, but such sampling programmes are time-consuming and expensive. Lohr (2010) stated that by far the best method of dealing with non-response is its prevention. In the context of this study, prevention would involve building trust with fishers or incentivising cooperation such that the risk of non-response would be low and less biased. There is a legal obligation to provide such data upon request, which if enforced would result in 100% response rate. However, our experiences in the two pilot studies highlighted some possible negative impacts of mandatory provision. Unreported catches are a sensitive issue and scientific results will have direct consequences to control and management of the fisheries. A willingness to contribute data can increase the acceptance of results (Hoare *et al.*, 2011; Mangi *et al.*, 2016), so fostering a trustful cooperation with the fishing industry at the earliest possible stage will improve the long-term success of monitoring of unreported catches.

If non-response cannot be prevented, then an understanding of the statistical properties of the non-respondents would reveal potential biases. Fortunately, daily logbooks, sales notes data and vessel GPS tracking provide highly detailed information on the characteristics of offshore vessels. Taking a representative sample of non-respondents would allow for inference about other non-respondents (Lohr, 2010).

Considerable efforts were taken by companies to provide production reports and by research staff to standardise and compile the data. Nevertheless, some reports covered the wrong time period and areas, or were summarised at the wrong resolution. The excessive time spent in processing data and rectifying issues suggests that the data collection methodology is not scalable to other case studies or over multiple years.

5.2. Comparison with total catch sampling to identify high-grading patterns

5.2.1. Information loss

We base our analysis on the assumption that any differences between size distributions of total and landed catches can be attributed to high-grading. However, that assumption breaks down as we must further process data before making a comparison. Firstly, the statistical methods may introduce biases which cannot be quantified, but which may affect the interpretation of results. Secondly, uncertainties introduced by the data processing methods will likely increase uncertainty, making it more difficult to identify small effect sizes that come with low levels of high-grading. This is an important consideration given that it is commonly assumed that Norway has low levels of discarding (Pérez Roda *et al.*, 2019; Gilman *et al.*, 2020) due to high compliance and a well-established discard ban.

Supplementing production report information with summarised data from the grading machine (average product weight) may alleviate the issue of estimating the number of fish caught (NOU, 2019), but it does not address the issues of coarse grading systems, censored size grades, and the possibility of multimodal distributions. The only way to alleviate these issues using production report data would be to use the individual fish measurements from the graders.

5.2.2. Converting product weight to length

The lack of observations of fish lengths in on-board factories requires a two-stage conversion before analysis. Official conversion factors for round weight by the Norwegian Directorate of Fisheries are presented for each product as single annual values with no measure of uncertainty. In reality, product yield varies depending on the quality of the cut and size of the fish, the latter of which varies both spatially and seasonally (Mello and Rose, 2005). Converting round weight to length or age is further necessary for results to be suitable for input into stock assessments. Both these conversions will introduce uncertainty (and possibly bias), which should be accounted for in final estimates of high-grading. Ignoring this uncertainty could risk a type I statistical error where significantly important levels of high-grading are reported due to misleadingly small uncertainties in the estimate.

Uncertainty in product weight to round weight conversions are available for selected species and products (Blom, 2014), whilst weight-length data are available from Norwegian Reference Fleet sampling to estimate the relationship. Both these sources of uncertainty can be factored into estimations using Monte Carlo simulations. This involves repeatedly generating random observations from within the known range of uncertainty to produce a full range of possible outcomes for the final estimate.

In the absence of sufficient information on uncertainty, a sensitivity study would help to understand the level of uncertainty necessary to cast doubt on the interpretation of high-grading estimates, and whether these potential levels of uncertainty are realistic.

5.3. Opportunities for future sampling

Production reports can be supplemented with data from the product grading system to provide the average weight (and therefore estimated number) of fish in the defined period. Data from individual fish passing over the grader are archived to generate a production report for any desired time period. Having direct observations of individual fish removes the need to infer an underlying size distribution, allowing for a direct comparison with size distributions of total catches by the Norwegian Reference Fleet. Chapter 6 of the Marine Resources Act 2008 gives a legal basis for a future implementation of a regulation that can allow for the acquisition of fish grading data from any fishing vessel for management or scientific purposes.

We have identified four issues regarding the use of individual fish measurements from grading machines. Firstly, there is the issue of cooperation with the fishing industry. We experienced some negative reactions from fishing companies to requests for aggregated data. Without addressing these issues, we would only expect concerns to increase if more detailed data were requested.

Secondly, the quantity of data generated by individual fish measurements would require a different solution for data transfer and storage. We have met with two leading companies that sell and maintain on-board fish grading systems to the Norwegian fishing industry. These companies can remotely access the system, and it is possible to transfer data across the connection. However, it is vital that agreements are developed between the Norwegian Directorate of Fisheries, fishing industry and grading technology companies to agree on a trusted and safe routine.

Thirdly, we need to develop a robust sampling design to ensure the reliability of estimates. For example, numerous studies in Norwegian offshore fisheries have stressed the importance of increasing the number of vessels (Helle and Pennington, 2004) and trips (Aanes and Pennington, 2003) sampled, demonstrating it is unnecessary to take large samples of fish from individual hauls (Pennington and Helle, 2011). A devoted pilot study would help to define the optimal sampling design, considering costs and excessive collection of sensitive data.

Finally, it is possible for fish to bypass the grading machine, and there is a risk that those fish bypassed may differ from those observed. We have no direct knowledge of when this occurs, but it is possible to infer by comparing grading data with production report summaries that include the bypassed fishes. Total weights in production reports are estimated, but if all fish were graded then the two totals should be equal within an acceptable degree of uncertainty.

Since 2019, the offshore segment of the Norwegian Reference Fleet began sampling discards, fishmeal and landed catches separately. High-grading could be identified by comparing the size distributions of these fractions, which could then be extrapolated to unsampled vessels to quantify the extent of high-grading in the fishery. Whilst this analysis could be done using only observations from the Norwegian Reference Fleet, we argue that individual size measurements from graders could supplement the analysis to include more vessels (Helle and Pennington, 2004).

Unfortunately, individual observations from the fish grader do not provide us with the ultimate need for length-based observations. However, they can be estimated using empirical length-weight relationships. A government report on the future of fisheries control (NOU, 2019) proposes an automatic documentation system where catches are registered at the earliest possible point after hauling. Such a system would provide direct observations of gross catches before any processing, removing the need estimate unreported catches, and therefore removing all the issues met in this report. For example, observations of fish before processing would remove the uncertainty surrounding the conversion from product weight to round weight. Likewise, laser measurement of length would also remove uncertainties surrounding a weight-length conversion. Any concerns would then be regarding the reliability of such a system, which would be dependent on enforcement strategy. However, a comparison with an equivalent, reliable source (e.g. gross catches recorded by the Norwegian Reference Fleet) could help to evaluate the reliability.

6. Acknowledgements

We are grateful for the cooperation of the fishing industry in both pilot studies. We also thank Vigdis Rong for the large task of standardising and compiling data. We finally thank David Hirst and Jon Helge Vølstad for discussions on the statistical analysis of the data.

7. References

Aanes, S., and Pennington, M. 2003. On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. ICES Journal of Marine Science, 60: 297–303.

Batsleer, J., Hamon, K. G., van Overzee, H. M. J., Rijnsdorp, A. D., and Poos, J. J. 2015. High-grading and over-quota discarding in mixed fisheries. Reviews in Fish Biology and Fisheries, 25: 715–736.

Berg, H. S. F., and Nedreaas, K. 2020. Estimering av utkast i Norsk kystfiske med garn 2012-2018. Institute of Marine Research.

Blom, G. 2014. A Norwegian guide for controlling bodies' inspection of the application of conversion factors on board fishing vessels. Directorate of Fisheries, 64 p.

Breivik, O. N., Storvik, G., and Nedreaas, K. 2017. Latent Gaussian models to predict historical bycatch in commercial fishery. Fisheries Research, 185: 62–72.

Clegg, T., and Williams, T. 2020. Monitoring bycatches in Norwegian fisheries - Species registered by the Norwegian Reference Fleet. Rapport fra havforskningen; 2020-8. https://hdl.handle.net/11250/2685855 (Accessed 10 June 2021).

Delignette-Muller, M. L., and Dutang, C. 2015. fitdistrplus: An R package for fitting distributions. Journal of Statistical Software, 64: 1–34.

Dickey-Collas, M., Pastoors, M. A., and van Keeken, O. A. 2007. Precisely wrong or vaguely right: simulations of noisy discard data and trends in fishing effort being included in the stock assessment of North Sea plaice. ICES Journal of Marine Science, 64: 1641–1649.

Dingsør, G. E. 2001. Estimation of discards in the commercial trawl fishery for cod. University of Bergen.

Gezelius, S. S. 2006. Monitoring fishing mortality: Compliance in Norwegian offshore fisheries. Marine Policy, 30: 462–469.

Gilman, E., Perez Roda, A., Huntington, T., Kennelly, S. J., Suuronen, P., Chaloupka, M., and Medley, P. A. H. 2020. Benchmarking global fisheries discards. Scientific Reports, 10: 1–8. Nature Publishing Group UK.

Gullestad, P., Blom, G., Bakke, G., and Bogstad, B. 2015. The “Discard Ban Package”: Experiences in efforts to improve the exploitation patterns in Norwegian fisheries. Marine Policy, 54: 1–9.

Helle, K., and Pennington, M. 2004. Survey design considerations for estimating the length composition of the commercial catch of some deep-water species in the northeast Atlantic. Fisheries Research, 70: 55–60.

Hirst, D., Aanes, S., Storvik, G., Huseby, R. B., and Tvete, I. F. 2004. Estimating catch at age from market sampling data by using a Bayesian hierarchical model. Journal of the Royal Statistical Society: Series C (Applied Statistics), 53: 1–14.

Hoare, D., Graham, N., and Schon, P. J. 2011. The Irish Sea data-enhancement project: comparison of self-sampling and national data-collection programmes-results and experiences. ICES Journal of Marine Science, 68: 1778–1784.

Huston, M., and Deangelis, D. 1987. Size Bimodality in Monospecific Populations: A Critical Review of Potential Mechanisms. American Naturalist - AMER NATURALIST, 129.

Lohr, S. L. 2010. Sampling: Design and Analysis. Chapman and Hall/CRC. 596 pp.

Mangi, S. C., Smith, S., and Catchpole, T. L. 2016. Assessing the capability and willingness of skippers towards fishing industry-led data collection. Ocean & Coastal Management, 134: 11–19.

McBride, M. M., and Fotland, Å. 1996. Estimation of unreported catch in a commercial trawl fishery. Journal of the Northwest Atlantic Fisheries Society, 18: 31–41.

Mello, L. G. S., and Rose, G. A. 2005. Seasonal cycles in weight and condition in Atlantic cod (*Gadus morhua* L.) in relation to fisheries. ICES Journal of Marine Science, 62: 1006–1015.

Methratta, E. T., and Link, J. S. 2007. Ontogenetic variation in habitat association for four groundfish species in the Gulf of Maine –Georges Bank region. Marine Ecology Progress Series, 338: 169–181.

Norwegian Directorate of Fisheries. 2021. Norske Omregningsfaktorer - For omregning av landet produktvekt av marin fisk og andre marine arter til rund vekt - for fiske i det nordlige Atlanterhavet og i andre farvann. https://www.fiskeridir.no/Yrkesfiske/Tema/Omregningsfaktorer (in Norwegian; Accessed 10 June 2021).

NOU. 2019. Framtidens fiskerikontroll. Departementenes sikkerhets- og serviceorganisasjon 2019:21 Oslo (in Norwegian).

Pálsson, Ó. K. 2003. A length-based analysis of haddock discards in Icelandic fisheries. Fisheries Research, 59: 437–446.

Pennington, M., and Helle, K. 2011. Evaluation of the design and efficiency of the Norwegian self-sampling purse-seine reference fleet. ICES Journal of Marine Science, 68: 1764–1768.

Pérez Roda, M. A., Gilman, E., Huntington, T., Kennelly, S. J., Suuronen, P., Chaloupka, M., and Medley, P. 2019. A third assessment of global marine fisheries discards. FAO Fisheries and Aquaculture Technical Paper No. 633. Rome, FAO. 78 pp.

Perretti, C. T., Deroba, J. J., and Legault, C. M. 2020. Simulation testing methods for estimating misreported catch in a state-space stock assessment model. ICES Journal of Marine Science, 77: 911–920.

Pitcher, T. J., Kalikoski, D., Pramod, G., and Short, K. 2009. Not honouring the code. Nature, 457: 658–659.

Punt, A. E., Smith, D. C., Tuck, G. N., and Methot, R. D. 2006. Including discard data in fisheries stock assessments: Two case studies from south-eastern Australia. Fisheries Research, 79: 239–250.

Reid, D. G., Calderwood, J., Afonso, P., Bourdaud, P., Fauconnet, L., González-Irusta, J. M., Mortensen, L. O., *et al.* 2019. The best way to reduce discards is by not catching them! *In* The European Landing Obligation: Reducing Discards in Complex, Multi-Species and Multi-Jurisdictional Fisheries, pp. 3–26. Springer, Cham.

Rizzi, S., Gampe, J., and Eilers, P. H. C. 2015. Efficient Estimation of Smooth Distributions from Coarsely Grouped Data. American Journal of Epidemiology, 182: 138–147.

Rizzi, S., Thinggaard, M., Engholm, G., Christensen, N., Johannesen, T. B., Vaupel, J. W., and Lindahl-Jacobsen, R. 2016. Comparison of non-parametric methods for ungrouping coarsely aggregated data. BMC Medical Research Methodology, 16: 1–12. BMC Medical Research Methodology.

Stenevik, E. K., Høines, Å., Kvamme, C., Otterå, H., Salthaug, A., and Svendsen, B. V. 2020. Fangstprøvelotteriet 2019 - Erfaringer og foreløpige resultat. Rapport fra Havforskningen nr. 2020-17 (in Norwegian).

Venables, W., and Ripley, B. 2002. Modern Applied Statistics with S. Springer, New York. https://www.stats.ox.ac.uk/pub/MASS4/ (Accessed 19 May 2021).

FISKERIDIREKTORATET

Telefon: 55 23 80 00
E-post: postmottak@fiskedir.no
Internett: www.fiskeridir.no

Livet i havet - vårt felles ansvar

Graphic design: Communication Division, UIB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230867624 (print)
9788230845868 (PDF)