BENCHMARK WORKSHOP ON PANDALUS STOCKS (WKPRAWN)

VOLUME 4 | ISSUE 20

ICES SCIENTIFIC REPORTS

RAPPORTS
SCIENTIFIQUES DU CIEM

ICES INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA
CIEM CONSEIL INTERNATIONAL POUR L'EXPLORATION DE LA MER

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

ISSN number: 2618-1371

ICES Scientific Reports

Volume 4 | Issue 20

BENCHMARK WORKSHOP ON PANDALUS STOCKS (WKPRAWN)

Recommended format for purpose of citation:

ICES. 2022. Benchmark workshop on Pandalus stocks (WKPRAWN).
ICES Scientific Reports. 4:20. 249 pp. http://doi.org/10.17895/ices.pub.19714204

Editors

Colm Lordan • Johan Lövgren

Authors

Michael Wall Andersen • Tanja Baagoe Buch • Sergey Bakanev • Krista Baker • Ewen D. Bell • Mikaela Bergenius Nord • Patrik Börjesson • Tanja Baagoe Buch • Anndorte Burmeister • Massimiliano Cardinale • José Miguel Casas Sanchez • Ole Ritzau Eigaard • Chris Griffiths • Carsten Hvingel Francesco Masnadi • Katja Norén • Cóilín Minto • Alessandro Orio • Mark Simpson • Katherine Skanes Guldborg Søvik • Fabian Zimmermann

ICES
CIEM
International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Contents

i Executive summary iii
ii Expert group information v
1 Introduction 1
2 Resolution 3
3 pra.27.3a4 North Sea in the Norwegian Deep 5
3.1 Stock ID and sub-stock structure 5
3.2 Issue list 5
3.3 Ecosystem drivers - predation index 11
3.4 Stock Assessment 13
3.4.1 Data 13
3.4.2 Correction factor 17
3.4.3 Survey 18
3.4.4 Biological information 23
3.4.5 Length-weight and growth 24
3.4.6 Maturity and size at sex change 24
3.4.7 Natural mortality 26
3.4.8 Assessment model 27
3.4.9 Model settings 29
3.4.10 The final model 36
3.4.11 Ensemble model 54
3.5 Short-term projections 61
3.6 MSE and Reference Points (MSY) 61
3.7 Future Research and data requirements 67
3.8 References 68
4 pra.27.12 Barents Sea (ICES Subareas 1 and 2) 71
4.1 Stock ID and sub-stock structure 71
4.2 Issue list 71
4.3 Ecosystem drivers 71
4.4 Stock Assessment 75
4.4.1 Catch – quality, misreporting, discards 75
4.5 Surveys 76
4.5.1 Barents Sea ecosystem survey 76
4.5.2 Barents Sea demersal fish survey in winter 79
4.6 Commercial CPUE standardization 83
4.7 Model priors 86
4.7.1 Carrying capacity (K) 86
4.7.2 Maximum Sustainable Yield (MSY) 88
4.7.3 Initial relative biomass (B_0/B_{msy}) 88
4.7.4 Observation and process error 88
4.7.5 Catchability or scaling parameters (q) 88
4.7.6 Shape parameter (n) 88
4.8 Length information and recruitment index 89
4.9 Assessment model 91
4.9.1 Comparison of assessment frameworks 91
4.9.2 Effects of revised time-series and priors 92
4.9.3 SPiCT stock assessment 93
4.9.4 Candidate models 93
4.9.5 Reference model 94
4.9.6 Model diagnostics and retrospective patterns 95
4.9.7 Sensitivity analysis 97

4.10 Short-term projections ... 99
4.11 Appropriate Reference Points ... 100
4.12 Future Research and data requirements ... 100
4.13 References ... 102
5 pra.27.3M Flemish Cap ... 104
5.1 Stock ID and sub-stock structure ... 104
5.2 Issue list ... 104
5.3 Ecosystem drivers ... 106
5.4 Fishery dependent data ... 106
5.5 Survey data ... 106
5.6 Assessment model ... 106
5.7 Future Research and data requirements ... 107
6 Reviewer reports ... 108
7 Recommendations ... 121
Annex 1: List of participants ... 122
Annex 2: Working documents ... 124
WD 1: Historic landings of northern shrimp (*Pandalus borealis*) in Norway. Guldborg Søvik. ... 124
WD 2: Fisheries data from the Norwegian commercial northern shrimp (Pandalus borealis) fishery in Skagerrak and the Norwegian Deep (ICES Divisions 3.a and 4.a East) ... 130
WD 3: Stock assessment of Northern shrimp (Pandalus borealis) in divisions 3.a and 4.a East (Skagerrak and Kattegat and northern North Sea in the Norwegian Deep) ... 157
WD 4 Natural Mortality for pra.27.3a4 ... 158
WD 5 Sampling design and estimation of commercial catches of northern shrimp Pandalus borealis by the Swedish bottom otter trawl fisheries in the 2016- 2020 ... 162
WD 6 Draft. Description of methodology behind compilation of Swedish data on Pandalus borealis (1908-2020) used in *Pandalus* benchmark (WKPRAWN 2022) ... 181
WD 7 Ensemble Model of Northern shrimp (Pandalus borealis) in ICES division 3a and 4a east ... 188
WD 8 Shrimp in the Barents Sea - assessment models ... 189
WD 9 Estimating size at sex change northern shrimp (*Pandalus borealis*). ... 190
WD10 Estimating a predation index for the 27.3a-4s northern shrimp (Pandalus borealis). ... 210
WD11 Estimating weight changes of processed shrimp catches in fisheries for northern shrimp (*Pandalus borealis*) in Skagerrak, Kattegat and the Norwegian Deep (ICES divisions 3.a and 4.a East) ... 224
WD12 Landings, discards, data coverage and quality from Denmark ... 244

i Executive summary

Three Northern shrimp (*Pandalus borealis*) stocks; pra.27.3a4a –divisions 3.a and 4.a East (Skagerrak and Kattegat and northern North Sea in the Norwegian Deep), pra.27.1-2 –in subareas 1 and 2 (Northeast Arctic) and the Flemish Cap (NAFO Div. 3M) stock were examined during this benchmark. The data and information available for all three stocks was very different and this largely dictated the assessment approaches investigated during the meeting.

There was and extensive issue list for pra.27.3a4 and ambitious plans to carry out the first ensemble benchmark assessment in ICES. Significant progress was made on the issue list with input landings, discard and biological data documented in excellent detail for all countries participating in the fishery. The only survey time-series was revised with a new index estimated with a latent spatio-temporal Gaussian model in StoX. Natural mortality was one of the key issues identified and four different M assumptions were tested during the benchmark. There was extensive work to estimate the size at sex change (L_{50}) it was decided that the Norwegian survey data were the most appropriate data to use for the L_{50} estimates, as the survey also targets younger, smaller individuals.

SS3 was chosen as the main assessment method and there was extensive model development work carried out in advance and presented at the meeting. This involved exploring various model and data configurations. During the course of the meeting additional data explorations and assessment runs were requested by both participants and reviewers, and the responsible scientists responded efficiently to these requests. The base model is a two-area, two-sex, quarterly age-length model that accounts for hermaphroditic reproductive life history of the prawns. The model was fit to include partial historical catches but given the similarity of most parameters, the most recent period were all catches were present was chosen (1970 onwards). An overall growth curve is estimated in the model. Recruitment is modelled as a Beverton-Holt function with a suitable value of the recruitment deviation variance set and parameters of the SR relationship estimated within the model.

An ensemble model was developed to incorporate uncertainty in natural mortality. A distribution of possible natural mortalities at age was constructed based on a suite of M life history and other type models. Three M scenarios were developed: a low, median and high M. An objective weighting approach was used based on model performance statistics and diagnostics to develop the final assessment. The reference points were estimated using a short cut MSE approach testing a range of fishing mortalities and biomass thresholds in relation to virgin stock size (B_0). B_{lim} for this stock was set at 15% of B_0 which was approximately the average B_{loss} for the three models (B_{loss} range: 11-17% of B_0). The group agreed a combination of an F_{MSY} proxy at $F_{30\%}$, combined with $B_{threshold}$ at 80% of $B_{30\%}$ satisfied the criterion of being above B_{lim} with >95% probability and generating catches within 95% of MSY.

For the Barents Sea stock, a bespoke Bayesian surplus production model has previously been used to assess the stock. There was no new information on stock structure and no clear signals were found in the analysis of shrimp biomass trends in relation to cod biomass and environmental indices. The commercial CPUE used in the assessment was thoroughly investigated and standardised using a non-spatial model. A spatiotemporal modelling approach (sdmTMB) for the joint Norwegian/Russian Barents Sea Ecosystem Survey (BESS) was presented and compared to the design-based estimators currently used. The modelled approach handled changes in coverage well and gave consistent estimates over time and was taken forward as an input for the new assessment.

The model exploration involved testing the BUGS and SPiCT surplus production models. The various priors were carefully considered with an informative prior for carrying capacity, K, was constructed based on a K-estimate for the West Greenland shrimp stock. The overall trends between both models were very similar and it was concluded that selection of any of the two methods of implementation would significantly influence assessment results. A number of candidate SPiCT models were explored in more detail and evaluated based on the criteria for the acceptance of a SPiCT model. The reference model included a prior on K and used a the commercial CPUE time-series. There was some retrospective error on F/F_{MSY} but Based on the presented diagnostics, the reference model was considered as adequate and acceptable. The standard approach for short-term forecasts and reference points were agreed.

There was limited progress on the Flemish Cap stock (pra.27.3M). An excellent long-term EU survey from 1988-present is available for the stock covering the period of the moratorium. This shows an increase 2016-2019 with a subsequent decrease. Length-composition data are also available from the survey for this assessment. In the future it may be possible to explore and SS3 model for this stock given the partial and patchy nature of the available data.

A number of priorities for future work were identified for all three stocks. For pra.27.3a4 improving the live weight correction, exploring time varying M, different weightings within the ensemble, exploring growth and length-based sex change were all suggested. For pra27.12 stock ID, cod predation, developing a recruitment index, estimating and including index uncertainty, combining surveys exploring seasonal dynamics and getting better fishery data were all suggested. For pra27.3M the key issue is to develop the stock assessment model.

ii Expert group information

Expert group name	Benchmark workshop on *Pandalus* stocks (WKPRAWN 2022)
Expert group cycle	Annual
Year cycle started	2022
Reporting year in cycle	1/1
Chairs	Colm Lordan, Ireland
	Johan Lövgren, Sweden
Meeting venues and dates	8–22 October 2021, Data Compilation WK, Online
	24–28 January 2022, Benchmark WK, Online

1 Introduction

WKPRAWN took place in Q4 2021 and Q1 2022 during the COVID pandemic which meant that the data evaluation workshop (DEW) and benchmark meetings were online. There were 36 participants from 8 countries including one participant representing the fishing industry. The two chairs were independent from Ireland and Sweden as were the two reviewers from Ireland and the United Kingdom. Not all members participated throughout and in reality, there were around 15 active participants. Johan Lövgren chaired the DEW and co-chaired parts of the benchmark meeting when Colm Lordan the chair of the Benchmark meeting was not available. Participants were asked to declare any conflict of interest (none were declared) and reminded of meeting etiquette at the start of each meeting.

The online sessions were good for information sharing and initial review but did not facilitate problem solving or collaboration across teams. Resources were also not well distributed across stocks. There was a big team working on pra.27.3a4a, 2-3 persons working on pra.27.1-2 and only one person working on *Pandalus* in 3M. The meeting was the first ever ICES benchmark to use an ensemble model so in that sense it was developing new approaches, methods and standards that have not been used elsewhere in ICES to date.

Twelve working documents were presented by participants and the contents of these were either included in the main report sections or have been appended to this report in Annex 2.

Table 2.1 Working Documents Presented to WKPRAWN 2022.

Working Document	Title	Author(s)	Report Section
WD1	Historic landings of northern shrimp (*Pandalus borealis*) in Norway	Katrine Wilhelmsen Melaa, Fabian Zimmermann and Guldborg Søvik	3.4.1
WD2	Fisheries data from the Norwegian commercial northern shrimp (*Pandalus borealis*) fishery in Skagerrak and the Norwegian Deep (ICES Divisions 3.a and 4.a East)	Guldborg Søvik, Trude H. Thangstad, Fabian Zimmermann	3.4.1
WD3	Stock assessment of Northern shrimp (*Pandalus borealis*) in divisions 3.a and 4.a East (Skagerrak and Kattegat and northern North Sea in the Norwegian Deep)	Massimiliano Cardinale, Alessandro Orio, Mikaela Bergenius Nord, Katja Norén and Francesco Masnadi	3.4.8
WD4	Natural Mortality for pra.27.3a4	Francesco Masnadi	3.4.7
WD5	Sampling design and estimation of commercial catches of northern shrimp *Pandalus borealis* by the Swedish bottom otter trawl fisheries in the 2016- 2020	Nuno Prista , Hongru Zhai , Katja Norén , Sofia Carlshamrea , Annica de Groote	3.4.1
WD6	Draft. Description of methodology behind compilation of Swedish data on *Pandalus borealis* (1908-2020) used in *Pandalus* benchmark (WKPRAWN 2022)	Katja Norén	3.4.1
WD7	Ensemble Model of Northern shrimp (*Pandalus borealis*) in ICES division 3.a and 4.a east	Francesco Masnadi, Massimiliano Cardinale, Alessandro Orio, Christopher Griffiths and Mikaela Bergenius Nord	3.4.11

Working Document	Title	Author(s)	Report Section
WD8	Shrimp in the Barents Sea - assessment models	Fabien Zimmermann, Carsten Hvingel	4.9, 4.10, 4.11.
WD9	Estimating size at sex change northern shrimp (*Pandalus borealis*).	Mikaela Bergenius Nord, Patrik Börjesson, Guldborg Søvik	3.4.6
WD10	Estimating a predation index for the 27.3a-4s northern shrimp (*Pandalus borealis*).	Patrik Börjesson, Mikaela Bergenius Nord, Guldborg Søvik	3.4.7
WD11	Estimating weight changes of processed shrimp catches in fisheries for northern shrimp (*Pandalus borealis*) in Skagerrak, Kattegat and the Norwegian Deep (ICES divisions 3.a and 4.a East)	Guldborg Søvik, Mikaela Bergenius Nord, Ole Ritzau Eigaard, Katja Norén, Mats Ulmestrand	3.4.2
WD12	Landings, discards, data coverage and quality from Denmark	Ole Ritzau Eigaard	3.4.1

2 Resolution

The **Benchmark workshop on *Pandalus* stocks** (WKPRAWN 2022), chaired by External Chair Colm Lordan*, Ireland, and ICES Chair Johan Lövgren*, Sweden, and attended by invited external experts Cóilín Minto (Ireland), Ewan Bell (UK), will be established and meet online for a five-day data compilation workshop 18–22 October 2021, and a five-day benchmark workshop 24–28 January 2022 to:

a) Evaluate the appropriateness of data and methods to determine stock status and investigate methods for short-term outlook taking agreed or proposed management plans into account for the stocks listed in the text table below. The evaluation shall include consideration of:
   i. Life-history data;
   ii. Fishery-dependent and fishery-independent data;
   iii. Further consideration of environmental drivers, multispecies information, and ecosystem impacts for stock dynamics in the assessments and outlook;
b) Agree and document the most appropriate method for evaluating stock status and (where applicable) short-term forecast and update the stock annex as appropriate. Knowledge of environmental drivers, including multispecies interactions, and ecosystem impacts should be integrated into the methodology. A full suite of diagnostics (regarding data, retrospective behaviour, model fit, etc.) should be examined as a whole to evaluate the appropriateness of any model developed and proposed for use in generating advice. If no analytical assessment method can be agreed, then an alternative method for providing advice (ideally one of the WKLIFE X[1] methods) should be put forward;
c) Re-examine and update (if necessary) MSY and PA reference points according to ICES guidelines (see: Technical Document on reference points[2]);
d) Draft stock annexes as part of the benchmark outcomes;
e) Develop recommendations for future improvements of the assessment methodology and data collection;
f) As part of the evaluation:
   i. Conduct a five-day data compilation workshop (DCWK). Stakeholders are invited to contribute data (including data from non-traditional sources) and to contribute to data preparation and evaluation of data quality. Data, particularly catch information, should be collated as far back in time as possible. As part of the data compilation workshop consider the quality of data including discard and estimates;
   ii. Following the DCWK, produce working documents to be reviewed during the benchmark workshop at least seven days before the workshop.

[1] https://doi.org/10.17895/ices.pub.5985

[2] https://doi.org/10.17895/ices.advice.7891

Stock or issue	**Assessment lead**
pra.27.3a4a – Northern shrimp (*Pandalus borealis*) in divisions 3.a and 4.a East (Skagerrak and Kattegat and northern North Sea in the Norwegian Deep)	Mikaela Bergenius (Sweden); Max Cardinale (Sweden)
pra.27.1-2 – Northern shrimp (*Pandalus borealis*) in subareas 1 and 2 (Northeast Arctic)	Carsten Hvingel (Norway); Fabian Zimmerman (Norway)
Northern shrimp (*Pandalus borealis*) on the Flemish Cap (NAFO Div. 3M)	José Miguel Casas (Spain)

WKPRAWN 2022 will report by18 February 2022 for the attention of ACOM.

3 pra.27.3a4 North Sea in the Norwegian Deep

3.1 Stock ID and sub-stock structure

The shrimp in ICES Division 27.3.a (Skagerrak and Kattegat) and the eastern part of Division 27.4.a (Norwegian Deep) are currently assessed as one stock and are exploited by Norway, Denmark, and Sweden. A single stock assumption is supported by the genetics work of Knutsen *et al.* (2015), who showed that the stock in the Skagerrak and Norwegian Deep (SKND) is comprised of one biological unit. Smaller, genetically different stocks were identified in some fjords along the Norwegian and Swedish Skagerrak coasts, but as the fishery on these shrimp units is comparatively small, these stocks are not treated separately in the assessment. The shrimp stock in the in the Gullmarsfjord (Swedish west coast) is likewise included in the assessment unit, but have some additional management regulations to the rest of the Skagerrak and Norwegian Deep stock. This work was presented at the previous benchmark (ICES, 2016) and was not discussed further.

Despite the presence of one biological unit, it was noted that differences in size distribution between the two areas are often recorded in the survey and commercial catches. Similarly, differences in growth between the Skagerrak and the Norwegian Deep as well as sex differences in growth were discussed. Both aspects of growth were recommended for further research. There also exists indication of larval drift between the two areas and movement of juveniles from Skagerrak into the Norwegian Deep, as well as a likely loss of larvae from the Norwegian Deep into areas farther north along the Norwegian coast, due to the strong northward flowing coastal current (unpublished results). To evaluate the stock structure was not a term of reference for this meeting and was only discussed in the context of a two-area stock assessment model. Thus, further research is advised on the incorporation of larval drift into the assessment.

3.2 Issue list

The issue list compiled for the meeting are detailed below in Table 3.2.1. An extra column 'Conclusions and outcomes' has been added to provide concluding remarks or outcomes for each issue.

Table 3.2.1. Northern shrimp in division 3.a and 4.a East. Issue list for the WKPAND benchmark meeting.

Issue	Problem/Aim	Work needed/possible direction of solution	Data needed to be able to do this: are these available/where should these come from?	External expertise needed at benchmark (type of expertise/proposed name)	Conclusions and outcomes
Landings data	Inclusion of historical catches	Compilation of input data by national data contacts by metier, area, and country (quarterly or yearly?)	Data already available at National level	No	Quarterly landings data for all countries were provided by area from 1908–2021. Landings were corrected for weight loss during boiling on a country-by-country basis – see section 3.4.1 for more details. Work on boiling correction factors and on correction for weight gain of raw shrimp stored on ice is also presented in WD 11.
Discards data and documentation of sampling schemes	Inclusion of historical catches and documentation of recent catches	Compilation of input data by national data contacts by metier, area, and country	Data already available at National level	No	Swedish discard data were provided by quarter and area for 2008-2020. Norwegian discard data were estimated for 3.a based on Danish discard rations from 2009-2016, and from 2017-2020 based on data from the Norwegian Reference fleet. Data from 4.a exists only from 2019-2020. Danish discard data was provided by quarter and area from 2013-2020. Information on the calculation of Danish discard data and sampling schemes is provided in section 3.4.1 and WD 12. Information on Swedish data is provided in section 3.4.1. and WD 6 and WD 5.
Evaluate the migration from SAS to StoX	Model comparison	Compare outputs	Data already available at National level	No	Survey indices (numbers-at-length) per area and year for the new time-series 2006-2021 (survey conducted in Q1) were estimated using a latent spatio-temporal Gaussian model (Breivik *et al.*, 2021). Survey indices (numbers-at-length) per area for 1984-2002 (Q4) and 2004-2005 (Q2) were estimated using the software StoX (Johnsen *et al.*, 2019). The survey in autumn 2003 was conducted with a different survey gear and the data are therefore not included in the assessment. Biomass indices per area and year were estimated for the entire time-series using the latent spatio-temporal Gaussian model.

Issue	Problem/Aim	Work needed/possible direction of solution	Data needed to be able to do this: are these available/where should these come from?	External expertise needed at benchmark (type of expertise/proposed name)	Conclusions and outcomes
Tuning series	Standardisation of survey time-series	Existing survey time-series needs to be standardised according to depth, trawling speed, bottom temperature, varying coverage, gear, time of day and season Options: Vast model (Thorson & Barnett 2017; Thorson 2019) Casper Berg model (Berg *et al.,* 2014) Olav N. Breivik (Breivik *et al.,* 2021)	Data already available at National level	No	Survey indices (numbers-at-length) per area and year for the new time-series 2006-2021 were estimated using a latent spatio-temporal Gaussian model (Breivik et al. 2021). The same model was used for estimating biomass indices per area and year for the entire time-series (1984-2021).
Discards	Estimate historical discards	Two options: (1) Let the SS3 model estimate discards (2) Estimate discards outside the model and incorporate into input data	Compilation of data is detailed in a previous issue	No	Option 1 was used whereby SS3 estimates historical discards
Biological Parameters	The natural morality (M) assumptions of the stock assessment model need to be better justified/explored	Sensitivity analysis of the assessment model to different M assumptions: A range of constant Ms Time-varying M based on predator change indexes or temperature change indexes Age-varying M	The following data are available from different sources: (1) Constant M values are theoretical and can be substantiated from the literature (2) Predator abundance data are available	No	Four different M assumptions were tested: (1) Age-varying M low (2) Age-varying M median (3) Age-varying M high (4) Predator index based on work detailed in WD 10 For more details, see section 3.4.8 and WD 10. The predator index was later removed from the ensemble and will require further research. Temperature change indexes of M and constant Ms were not tested.

Issue	Problem/Aim	Work needed/possible direction of solution	Data needed to be able to do this: are these available/where should these come from?	External expertise needed at benchmark (type of expertise/proposed name)	Conclusions and outcomes
			from the survey. Temperature data can be obtained from ICES (3) Age-varying M assumptions are theoretical but can be substantiated from the literature		
Biological Parameters	Sex Ratio	Reconstruct sex ratio by year-at-length	Data already available at National level	No	During the data workshop, estimates of size at sex change (L_{50}) were presented based on Swedish commercial catch and Norwegian survey data. See WD 09. No apparent trends were visible in the data, and it was decided to: Use the Norwegian survey for the estimates Use a constant estimate of L_{50} in the assessment model It was decided that size at sex change = size at maturity. It was pointed out at the data workshop that small individuals <~ 12 mm are not sexed in the Norwegian survey. Moreover, there have been observations that 10-15% of small individuals might be primary females, thus L_{50} and SSB may not be entirely correct. To account for the presence of primary females in the commercial data, further research will investigate the consequences for estimates of L_{50} and SSB. It was also decided that further research will estimate L_{50} separately for 4.a East and 3.a, if possible, as growth between these two areas is expected to be different.
Assessment method	Develop a fleet-area based model in SS3	Compilation of input data by national data contacts	Data already available Input data available	No	A two-area based model was developed and presented to the meeting. The two areas represent the Skagerrak and Norwegian Deep (4.a East) respectively.

Issue	Problem/Aim	Work needed/possible direction of solution	Data needed to be able to do this: are these available/where should these come from?	External expertise needed at benchmark (type of expertise/proposed name)	Conclusions and outcomes
	Develop a hermaphroditic model Introduce an ensemble model that can incorporate several different model configurations that have similar diagnostics but different stock status results	Ensemble approach and software has been developed and is supported by ICES	for contemporary data but catches by sex need to be recalculated. Survey data by sex will be available with new Stox-index Assessment model data are already available		Landings, discards, and length frequency distributions were separated by area. Growth and other life-history traits (natural mortality and maturity, etc) were assumed to be fixed across areas and time. A hermaphroditic model was developed and presented to the meeting. Males are assumed to change to females according to a logistic curve, whereby all males have become females at age 2+. The inclusion of hermaphroditism greatly improved model fit, see section 3.4.9 and WD 03 for more details The hermaphroditic and area based SS3 model was accepted as the base case for the assessment. The base case is summarised in section 3.4.10 and detailed in WD 03. An ensemble model was developed and presented to the meeting (WD07). The ensemble considered 3 different model configurations, each with a different assumption for M (see above). Model results were aligned well with the base case and the ensemble approach was accepted by the meeting.
Biological Reference Points	Sensitivity of assessment results to the choice of B_{lim} Relative instead of absolute reference points	Sensitivity analysis Approach and software have been developed and is supported by ICES	Different B_{lim} assumptions will be tested Assessment model data are already available	No	Discussions on the value of B_{lim} were had during the meeting. It was decided that B_{lim} would be defined as a fraction of B_0. B_{lim} was set at 15% of B_0, which is approximately the average B_{loss} from the 3 models considered in the ensemble (B_{loss} range: 11-17% of B_0). For more details see section 3.6. The use of relative instead of absolute reference points is necessitated by the use of an ensemble model. This is because the scale is different between models, but the ratio remains constant. Relative reference points were presented and accepted by the meeting.
Forecast model	Currently the forecast is produced outside of the assessment model in	Integrate forecast within updated SS3 assessment model (available	Assessment model data are already available	No	The forecast is now integrated in the SS3 model. For more details see section 3.5.

Issue	Problem/Aim	Work needed/possible direction of solution	Data needed to be able to do this: are these available/where should these come from?	External expertise needed at benchmark (type of expertise/proposed name)	Conclusions and outcomes
	a complementary model	in newest version of SS3)			
Management strategy evaluation	Evaluation of Harvest Control Rules (HCRs) within the assessment model. Following the integration of the ensemble model and relative reference points, HCRs can be evaluated without external software/models	Approach and software tailored for the SS3 assessment model has been developed	Assessment model data are already available	Henning Winker from JRC joined the meeting	A MSE 'short-cut' approach was used in conjunction with developed methods for SS3. The MSE concluded that a $F_{B30\%}$ with $B_{trigger}$ set at 0.8 of $B_{30\%}$ is appropriate for the stock and will be employed in the assessment. For more details see section 3.6 (WD07).

3.3 Ecosystem drivers - predation index

In terms of ecosystem drivers of the *Pandalus* stock in the Skagerrak and Norwegian Deep area, work was conducted to create a natural mortality index based on predator abundance. This work is detailed in full in the WD 10 and was presented to the meeting. In brief, this work involved creating biomass and predation indices (1984-2021) for two assemblages of *Pandalus* predators in the Skagerrak and the Norwegian Deep. The first assemblage contained four species (roundnose grenadier, velvet belly, blue whiting, and Atlantic cod), each of which has previously been identified as primary predators on *Pandalus* (Jørgensen *et al.*, 2014). The second added starry ray, hake, haddock, whiting and saithe to the four primary predators. These two assemblages are referred to as the ‘4 species’ and ‘9 species’ index in the following text. Data were taken from the Norwegian bottom trawl survey for northern shrimp (NO-SH) and from the international bottom trawl survey in the North Sea (NS-IBTS Q1/Q3).

To calculate biomass indices, standardized catch per unit effort (CPUE) data was modelled on a species-by-species basis using a delta-GAM approach. Estimated CPUE’s were then combined by summing the estimated single species CPUE’s for either assemblage. In comparison, predation indices were calculated from the single species CPUE’s by applying species specific weights based on *Pandalus* consumption rates prior to aggregation. Combined indices were also calculated directly from the index of predator biomass that is annually taken during the Norwegian shrimp survey for comparison.

Direct comparison of biomass and predation indices show that the choice of predator composition (4 species vs. 9 species) has an appreciable impact on the index (figures 3.3.1 and 3.3.2), but that the inclusion of weights contributes little to the overall variation, i.e. there is very little difference between biomass and predation indices. One caveat here is that fixed weights were used throughout the entire period due to data constraints. This is a limitation because we fully expect that prey preferences will vary through time in response to changing stock size, the predator’s size distribution and the availability of other, potentially preferred prey items.

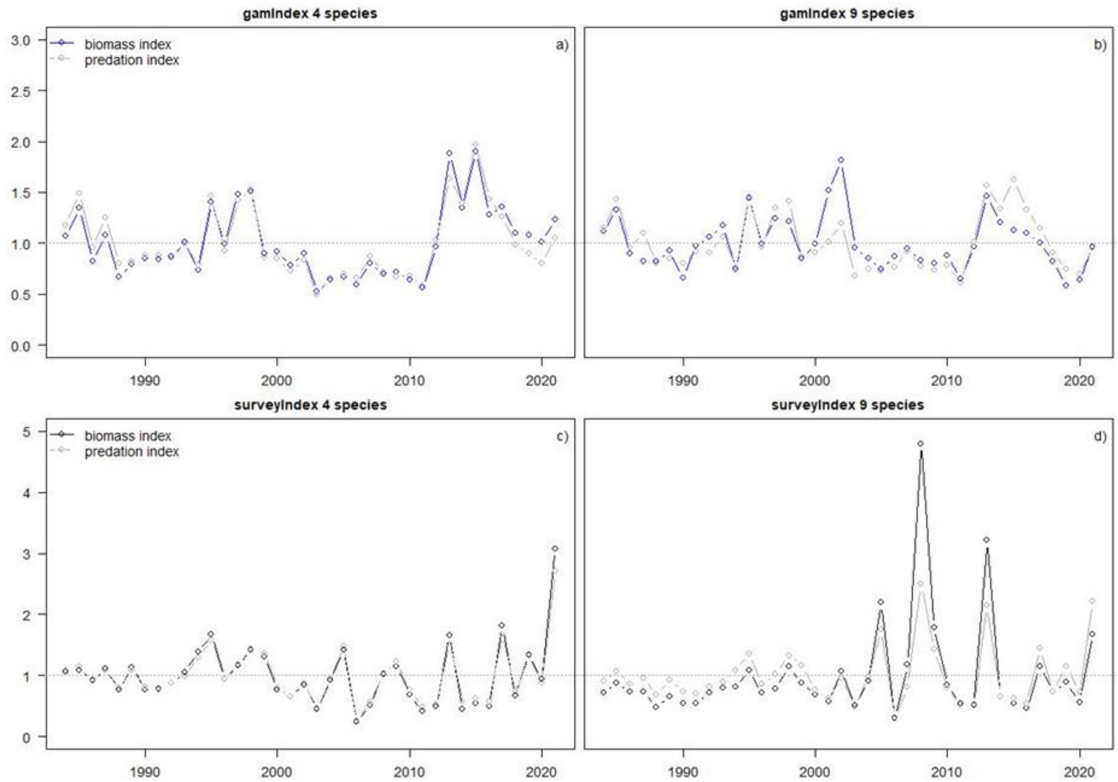


Figure 3.3.1. Comparison of the biomass and predation indices from 1984-2021 based on a) GAM model using 4 species, b) GAM model using 9 species, c) survey index using 4 species, and d) survey index using 9 species. Survey data for 2003 and 2016 were not available for modelling but were present in the survey index. Note the scale difference between the upper and lower panel.

The 4 species index and the nine species index showed similar trends, but with some notable differences. For example, in 2001 and 2002, the nine species index reached its highest values, whereas the four species index appears to be decreasing indicating a lower-than-average predator biomass. From 2011 to 2013 both indices showed a step increase, but while the nine species index started to decline immediately after, the four species index peaked again in 2015. These differences between the two indices are probably linked to species composition. In 2001 and 2002 the CPUE index for haddock and whiting peaked, and the CPUE of saithe was also going up. None of these species are included in the 4 species index and consequently it is unsurprising that the indexes behave differently. Further, in 2015 and 2016, the cod biomass reached its highest values in the time-series which is reflected in the four species index where cod is the only gadoid. In comparison, in the nine species index the peak in cod biomass is suppressed by the presence of other gadoids, which when coupled with a very low CPUE for roundnose grenadier, leads to a declining trend. It is also worth noting that the three deep sea species (roundnose grenadier, velvet belly and blue whiting) used in both indices are mainly caught at depths below 250 meters which is the depth limit for the NS-IBTS. In practice this means that the four species index is a modelled shrimp survey index, with added data on Atlantic cod from the NS-IBTS.

This modelling exercise illustrates the utility of combining data from the various surveys to generate biomass or predation indices for the Norwegian Deep / Skagerrak *Pandalus* population (NDSK). However, further work is required before the approach can be implemented in assessment. First, an evaluation of uncertainties in the model is needed. Second, model simplification and data pruning should be evaluated to improve model fit and prediction. Finally, the composition of predators in the index needs to be settled.

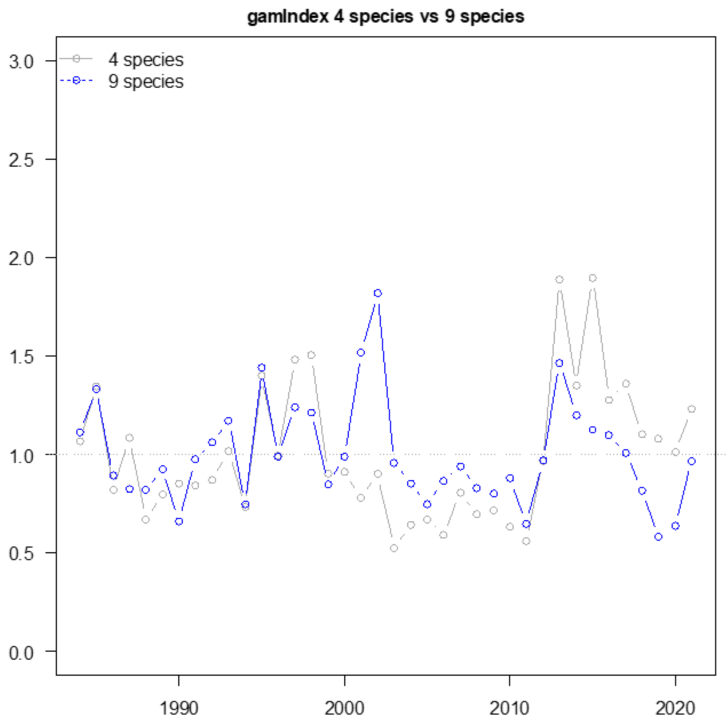


Figure 3.3.2. Comparison of the modelled biomass indices 1984-2021. Survey data for 2003 and 2016 were not available for modelling but were present in the survey index.

3.4 Stock Assessment

3.4.1 Data

General presentations of the available data sources were given from each country at the data evaluation workshop meeting (DEW). Below we outline fishery-dependent (landings and discards), fishery-independent, and biological data that are used as input data in the different assessment models.

Landings, discards, data coverage and quality from Sweden

In the benchmark, Swedish landings data for the period 1908-2020 were used. In earlier assessments, landings data were produced for the North Sea and Skagerrak/Kattegat combined but for this benchmark landings data per year, quarter and area (North Sea and Skagerrak/Kattegat separately) were requested. The methodology for producing landings per year, quarter and area is described in WD06 (Annex 2). Shrimps lose weight during boiling and therefore the weight of boiled shrimps has been accounted for by multiplying by a factor of 1.13 back to 1963. The procedure used for calculating boiled weight of *Pandalus* is described in the WD05 (Annex2).

Swedish discard information is derived from onboard sampling and self-sampling since 2008 as described in WD06 (Annex 2). Generally, onboard sampling is done on shrimp trawls with grid OTB_CRU_32-69_0_0 and by self-sampling on shrimp trawls with grid but no tunnel OTB_CRU_32-69_2_22. In earlier assessments discard values were calculated for the North Sea and Skagerrak/Kattegat combined but for this benchmark discard data was requested per year, quarter, and area (North Sea and Skagerrak/Kattegat separately). Discard values for 2016-2020 were recalculated using probability-based estimation as described in WD05 (Annex 2). Discard values for 2008-2015 were derived using historic values for the North Sea and Skagerrak/Kattegat combined and these were allocated proportionally according to information on landings per year, quarter and area as described in WD05 (Annex2).

The discard information derived from onboard sampling is regarded to be of better quality than the data derived from self-sampling. However, data from onboard sampling might be influenced by observer effects.

Landings, discards, data coverage and quality from Denmark

From 1988 to 2020, the benchmark data of DK landings from 27.3a4a is based on EU logbooks and Danish sales slips, which in combination deliver high-confidence full coverage landings data by year, quarter, and area (3.a and 4.a East). Prior to 1988, the benchmark data of DK catches (back to 1940) has been based on ICES statistics (https://www.ices.dk/data/Pages/default.aspx) and for this period a number of assumptions and extrapolations have been necessary to meet the desired spatial and temporal resolution (WD 12, Annex 2), and consequently the data from this source is associated with larger uncertainty. Danish discard data for 2009 to 2020 was included in the benchmark. From 2013 to 2020 the data is based on samples from 65 trips and 231 hauls of the Danish observer program (WD12). This corresponds to roughly 1-2% of all trips targeting *Pandalus borealis*. The discard estimates for this period are based on three separate sample-fractions per haul; 1) small, discarded shrimp, 2) medium-sized shrimp, landed raw, and 3) large shrimp, boiled on-board before being landed. The individual weights of the approx. 5-7 quarterly samples were summed by catch-fraction (1,2, and 3 above) before calculating a quarterly discard to landings ratio (the pooled weight of the discarded shrimp fraction to the pooled weight of the two landed fractions). Prior to the calculation of quarterly discard percentages, the total weight of fraction 3 (large, boiled shrimp) is multiplied with a factor of 1.13 to correct for weight loss. For 2009-2012 the discard data were taken from the ICES summary sheet and split by quarter and area using the average proportions as estimated with the higher-resolution data from 2013-2015 (3-years average). Finally, the quarterly discard ratios were applied to the total official landings (after correction for weight-loss due to on-board boiling and cooling) to provide total quarterly Danish discards of *Pandalus borealis* in 3a and 4a-east in weight for the years from 2009-2020.

As the Danish observer-based sampling data are almost exclusively from 3.a (only 5 of 65 trips are from .4a-east), the discard samples and data have been pooled for the two areas and the same quarterly discard ratios have been applied to the total landings from both areas.

Landings, discards, data coverage and quality from Norway

Historic landings

The Norwegian fishery for northern shrimp started in the late 1890s, as the large stocks of shrimp in the Oslofjord and eastern Skagerrak region were discovered through fishery investigations (Hjort and Ruud, 1938). Official landings statistics are available from 1908. Data from 1908 to 1976 were digitized in 2021. The historical data are given by county, which were separated into the ICES divisions 3.a and 4.a East. The following counties were placed in division 3.a: Viken, Oslo, Vestfold og Telemark, and Agder. Division 4.a includes Rogaland and Vestland. The county borders do not completely overlap with the borders of the ICES divisions (Figure 3.4.1).

There was a gradual increase in the Norwegian shrimp landings from 1908 to 1935, followed by a drop in landings during World War II (Figure 3.4.2). From 1950 to 1963, there was a steep increase in landings, followed by a large decline towards 1970. Landings then increased over the next 35 years from 1970 to 2005, and then gradually fell towards 2021.

Figure 3.4.1. Norwegian counties (thick blue line) and the Norwegian statistical grid (thin red line). Statistical area 09 corresponds to ICES Division 3.a and is defined to consists of the following counties: Viken, Oslo, Vestfold og Telemark and Agder, while area 08+28, corresponding to 4.a East is defined to consists of Rogaland and Vestland.

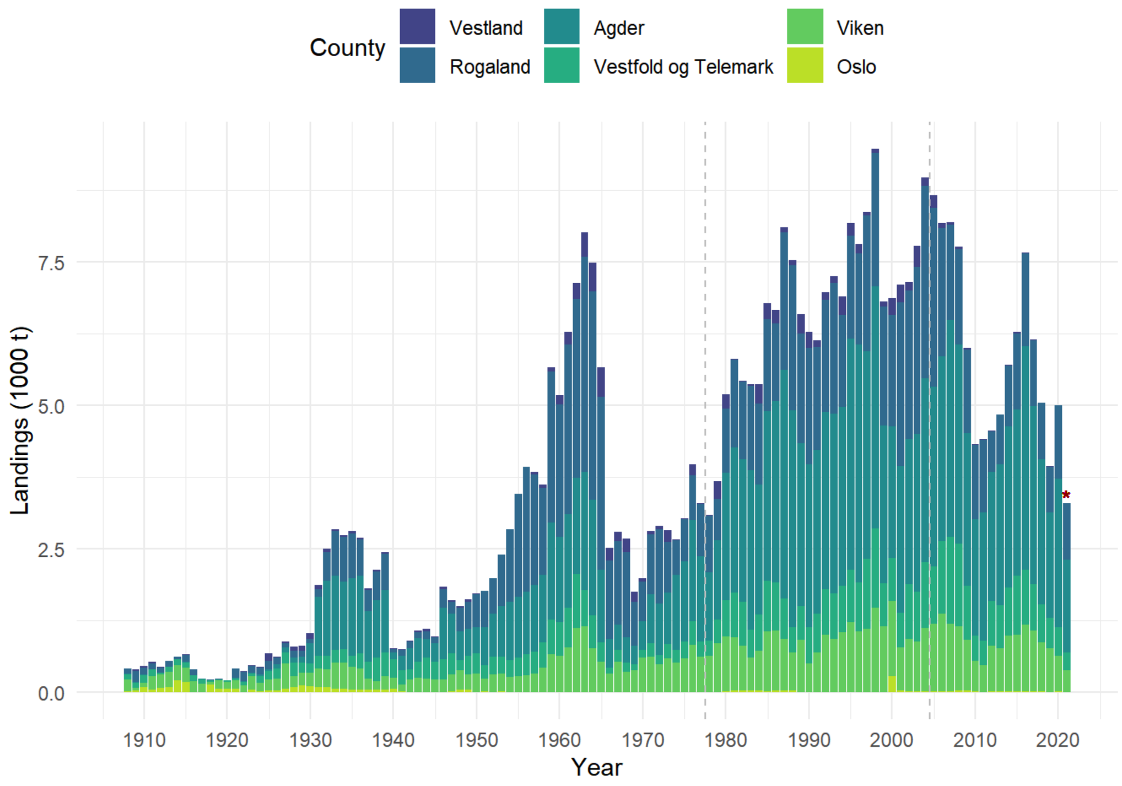


Figure 3.4.2. Historic Norwegian landings of northern shrimp (1000 t) by year and county in Skagerrak and the Norwegian Deep. Values for 2021 (*) include data until October.

Norwegian landings

Official Norwegian shrimp landings per year, quarter and area (divisions 3.a and 4.a) were obtained from the Norwegian Directorate of Fisheries, where information on landings per vessel trip has been available since 2005. Data from before 2005 are more aggregated.

Fishers sort the catch on board into three size fractions (sometimes two): large shrimp, medium-sized shrimp, and juvenile shrimp. The third size fraction may also contain glass shrimp (*Pasiphaea* sp.) and is usually discarded, but may also be landed. The medium-sized shrimp are landed raw, while the large females are boiled on board to be landed fresh, fetching high prizes. The official weight of the boiled landings has been corrected by multiplying by a factor of 1.13 to obtain live weight. Information on conservation of the Norwegian landings (boiled, raw) has been available since 2000, and boiled landings back to 2000 have been corrected.

Norwegian discards

Norwegian discards have been estimated since 2009. Discards were recalculated for the benchmark and were only estimated for areas and quarters with data available (WD01, Annex2). For the years 2009-2016, quarterly discards were estimated by applying the Danish discard ratio per quarter to Norwegian landings corrected for loss of weight due to boiling, assuming the same discard practice in the two countries. Data on Danish discarding in 2009-2016 only exist from Skagerrak, and Norwegian discards from the same years could only be estimated for this area.

Since 2017, discards have been estimated using data from the Norwegian Coastal Reference fleet (CRF), which consists of commercial vessels engaged by IMR to register and report all catches including discards (Hatlebrekke *et al.*, 2021). Since 2016, shrimp trawlers in the Skagerrak and Norwegian Deep have been included in this fleet. From all hauls, the vessels report weight of the three shrimp catch fractions (large, medium-sized, and juveniles/glass shrimp), and note whether the shrimp are landed or discarded. From approximately one haul every second week, vessels send samples from the three catch fractions to IMR where the samples are weighed, and length measured (up to 100 specimens). Discard samples are sorted into northern shrimp and glass shrimp, enabling the total discard weight to be partitioned into separate weights for northern shrimp and glass shrimp. Quarterly discard rate is estimated as a weighted average over all haul-wise discard rates. Due to few vessels in the CRF, an uneven distribution of hauls over division per quarter and year, and a limited number of trips and samples, discard rates were estimated for 3.a and 4.a combined for the years 2019-2020. In 2017-2018, there were no CRF vessels fishing in 4.a, so discard rates were not estimated for this area in these years. Quarterly discards were estimated by applying discard rates to official landings corrected for boiling.

3.4.2 Correction factor

In the shrimp fishery in the Skagerrak and Norwegian Deep, Norwegian and Swedish fishers traditionally have sorted the catch onboard, and the catch fraction consisting of the larger females has been boiled and landed fresh, fetching high prices. The Danish shrimp fishers started on board boiling of shrimp in the early 2000s and land the boiled catch primarily in Sweden. As shrimp lose weight when boiled, the shrimp working group (NIPAG) has corrected the commercial landings by multiplying the boiled catch fraction with a factor of 1.13, to obtain live weight. Norwegian and Danish landings have been corrected back to 2000 and 2001, respectively, and Swedish landings back to 1970.

Norwegian shrimp fishers have pointed out that a large fraction of the Norwegian fleet cool the boiled shrimp with water, not in air, and that cooling with water does not lead to weight loss. To investigate the effect of different cooling methods on weight loss due to boiling, an experiment was carried out in 2017 during the annual Norwegian shrimp cruise in the Skagerrak and Norwegian Deep (WD11). The results showed that cooling method does affect weight loss (Table 3.4.1).

Table 3.4.1. Correction factors based on boiling experiment on R/V Kristine Bonnevie in 2017, mean with SD, where n = 14 for water-cooled samples, and n = 13 for air-cooled samples.

	Factor_boiled	Factor_boiled-frozen	Factor_boiled-frozen-thawed
Water	1.00 (0.01)	1.00 (0.01)	1.01 (0.01)
Air	1.07 (0.02)	1.07 (0.04)	1.08 (0.02)

In Denmark, studies have been carried out to investigate if storing raw shrimp on ice affects the catch weight. This is the usual way of storing the catch at sea. Six trials carried out on board commercial vessels in 2020 and 2021 for up to six days, showed that the shrimp gained 7 % weight on average the first day on ice. In the following days, the shrimp added 1.4 % weight on average (Table 3.4.1) (WD11).

Table 3.4.2. Weight changes (%) with time from six different experiments with storing raw shrimp on ice.

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
nov-20	8.3	8.6	9.3	12.8	13.2	14.6
nov-20	6.5	8.5	9.1	9.7		
okt-20	5.3	9.0	8.1	10.3		
sep-21	8.5	10.7	14.4	12.0		
sep-21	7.2	11.6	11.6	12.5		
nov-21	5.9	8.6	9.9	10.0	11.4	
Average (%)	**7.0**	**9.5**	**10.4**	**11.2**	**12.3**	**14.6**
SD	1.3	1.3	2.3	1.4		

Preliminary conclusions are that the landings weight from the two catch fractions (raw and boiled shrimp) should be corrected to obtain live weight by downscaling the weight of the raw shrimp landings and upscaling the weight of the boiled shrimp landings. As new experiments are planned, NIPAG will await these before final conclusions are made regarding the most appropriate methods for correcting landings of shrimp to reflect live weight.

3.4.3 Survey

General survey description

Since 1984, a trawl survey for northern shrimp in Skagerrak and the Norwegian Deep has been conducted annually by the Norwegian Institute of Marine Research (IMR) with the objective of assessing the distribution, biomass, abundance, recruitment, size distribution, and demographic composition of the shrimp stock, the size of stocks of shrimp predators, as well as measuring hydrographical conditions in the area.

The survey data consist of: 1) one time-series from 1984 to 2002 (October/November) using R/V *Michael Sars* and the Campelen-trawl; 2) a point estimate for 2003 (October) as R/V *Michael Sars* was taken out of service and substituted with R/V *Håkon Mosby*, whose winches at that time were not powerful enough for the Campelen-trawl, resulting in the survey being conducted with the Åkra Shrimp trawl 1420; 3) a start of a potential new time-series as the survey in both 2004 and 2005 was conducted in May/June with R/V *Håkon Mosby* using the standard Campelen trawl; and 4) one time-series from 2006 until present (January/February), using R/V *Håkon Mosby* and R/V *Kristine Bonnevie* (from 2017) and the Campelen trawl. Conducting the survey in quarter 1 gives good estimates of the 1-group (recruitment) and SSB (berried females) and was recommended by the *Pandalus* working group in 2004 (ICES 2005).

The survey area covers depths of approximately 100 to 550 m. The survey is stratified by four depth zones (100-200 m, 200-300 m, 300-500 m, and >500 m), and area (Figure 3.4.1). The border between Skagerrak and the Norwegian Deep as defined in the strata system does not correspond with the present border between Divs. 3.a, and 4.a (a straight line between Hanstholm and Lindesnes).

The survey has a fixed station design with 111 stations (Figure 3.4.1). In some years, part of the survey area has not been covered due to time and weather constraints, most often this pertains to Division 4.a East as it is more exposed than Division 3.a.

A Campelen 1800/35 bottom trawl with rockhopper gear is used. Strapping was introduced in 2008 to ensure fixed trawl geometry. Mesh size in the cod end is 20 mm with an inner lining net (10 mm). In 2003, a Åkra Shrimp trawl was used (ICES 2005). Due to the different survey gear, data from 2003 has not been analysed. Tow duration was 1 hour until 1989 when it was reduced to 0.5 hours. Trawling is carried out around the clock. Tow speed is roughly 3 knots. Details are found in Søvik and Thangstad (2021).

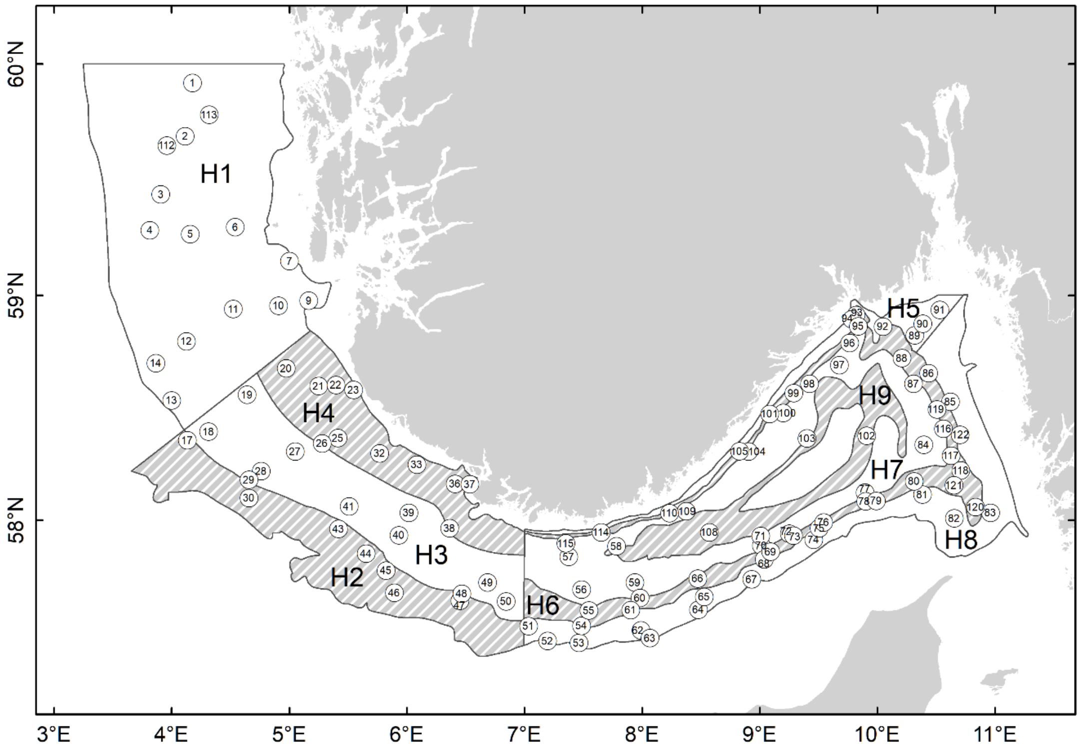


Figure 3.4.1. The strata system of the Norwegian bottom trawl survey in Skagerrak and the Norwegian Deep used for generating StoX indices, with positions of fixed trawl stations.

Stock indices

Until 2021, a design-based survey index estimated in SAS was used as input to the assessment model, calculating the abundance per length group as mean density raised to the corresponding area and depth stratum within the survey strata system (an older strata system than the one shown in Figure 3.4.1 was used, with a total of 17 strata). The design-based index has been re-estimated in the software StoX (Johnsen *et al.*, 2019). However, inconsistencies within the time-series in coverage due to weather and technical issues required *ad hoc* corrections for missing strata and resulted in the removal of the entire 2016 survey from the previous survey index. In addition, the design-based approach was found sensitive to the stratification because of the narrow depth contours, as well as start- and stop-locations of trawl stations that crossed in some cases strata or area borders. Model-based survey standardization that combines fixed effects such as bottom depth with random effects, including information from spatio-temporal correlation, has been shown to resolve such issues. The model applied here by Breivik *et al.* (2021) was specifically developed with the goal to improve the prediction of length-dependent abundance by using spatial random fields and correlation between length groups.

The model was fitted to abundance-at-length data from the 2006–2021 segment of the survey time-series. Based on model validation and testing, the configuration was tuned towards a sufficiently resolved spatial mesh, sun height was excluded as covariate due to lack of a significant effect, and length group intercepts were linked through a random walk process to improve estimates for specific length groups with few or no samples in some years. Total abundance and associated uncertainty were predicted by area. The resulting index followed largely the same trend as the design-based estimate with some notable deviations (Figure), particularly in years with coverage issues. Generally, the trend of the modelled index is slightly more pessimistic, although the design-based index remains within the 95% confidence intervals of the modelled index. Differences in specific years become more evident in the length frequency distributions (Figure).

While data from the 2016 survey was not included in the log-likelihood estimation and, thus, had no influence on the overall parameter estimates, abundance-at-length and uncertainty could be predicted from the model using the estimated fixed effects (bottom depth) and correlation structure. In contrast to the design-based index, the model-based approach enabled the estimation of an index for 2016 that adequately reflects the associated uncertainty and, thus, provides better information of the abundance in 2016 than excluding the year entirely. Following this significant methodological change, the re-inclusion of 2016 in the survey index for the stock assessment was accepted by WKPRAWN.

In addition to the length-based abundance index, the stock assessment also includes a total biomass index. Previously, the assessment model used total numbers. To ensure consistency in methods, total biomass was estimated using a spatio-temporal model that used the same model setup and configuration as the length-based index, simply using total observed catch weight as response instead of abundance-at-length. The model was implemented in the R package sdmTMB (Anderson *et. al*., 2020), a modelling framework specifically developed for the standardization of indices by including spatio-temporal correlation. Comparing the split estimates of old/new survey data with an estimate that uses the joint time-series showed that a joint biomass index may be preferable, because the mean estimates are very similar (Figure 3.4.4) while the longer time-series helps with the parameter estimation, resulting in lower uncertainty especially in the new time-series.

The stock indices for the two earlier time-series, 1984-2002 and 2004-2005, are still design-based indices. Survey coverage was less of an issue when the survey was carried out in October/November, but in 2002, the northernmost stratum H1 was not covered. Length frequency distributions show that the annual length-based abundance indices estimated by SAS and StoX are more or less identical except for 1997 (Figure 3.4.5).

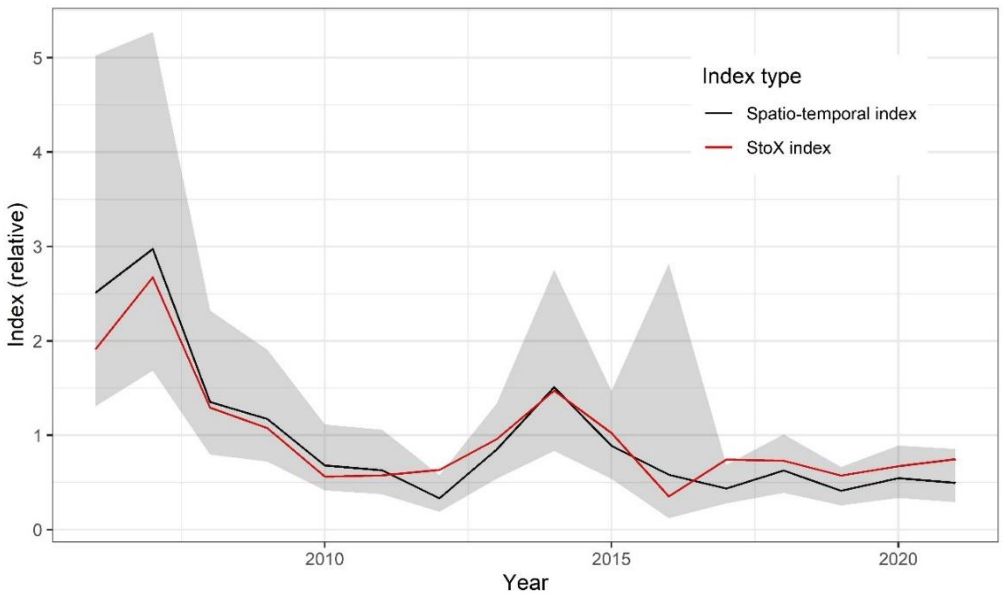


Figure 3.4.2. Abundance indices estimated through model-based (spatio-temporal index, black) and design-based (StoX index, red) approaches, shown as annual total abundance (sum of abundance-at-length). The grey area represents the 95% confidence intervals estimated in the model-based approach. Both indices were standardized to its mean.

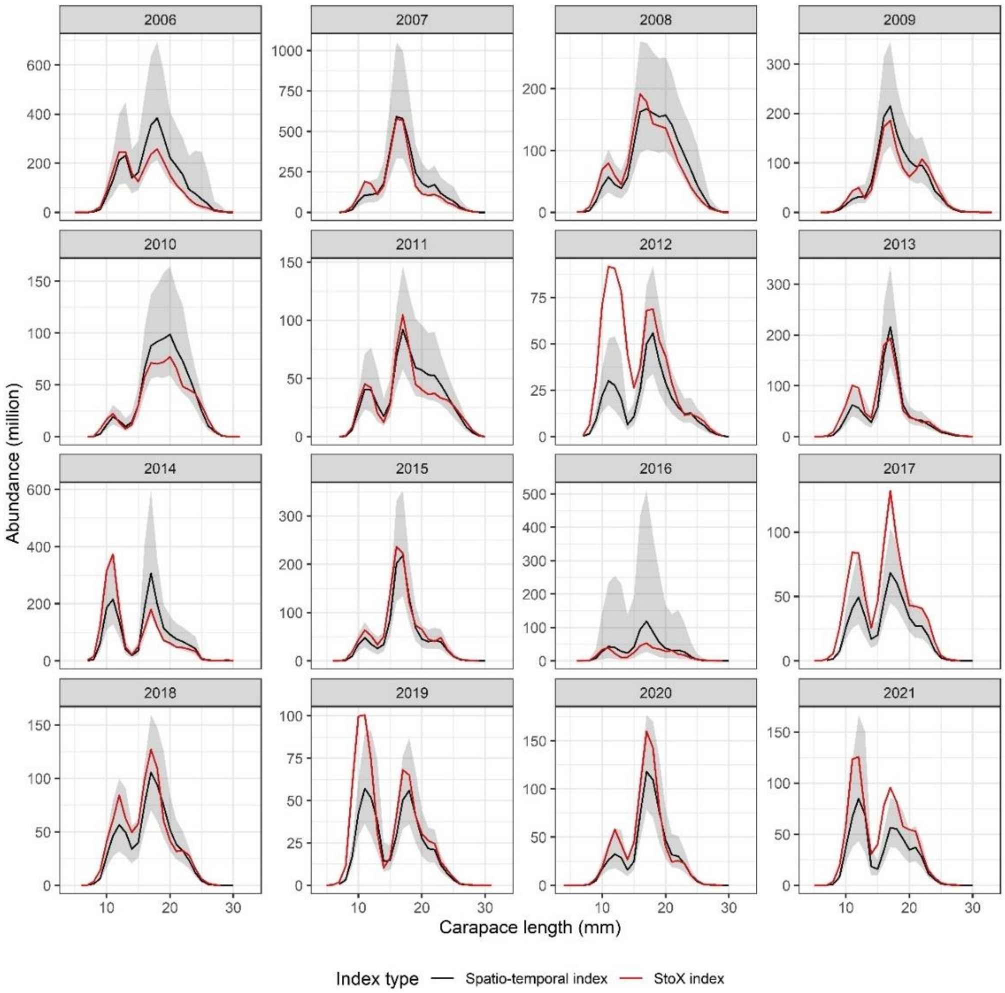


Figure 3.4.3. Abundance-at-length by year estimated through model-based (spatio-temporal index, black) and design-based (StoX index, red) approaches). The grey areas represent the 95% confidence intervals estimated in the model-based approach.

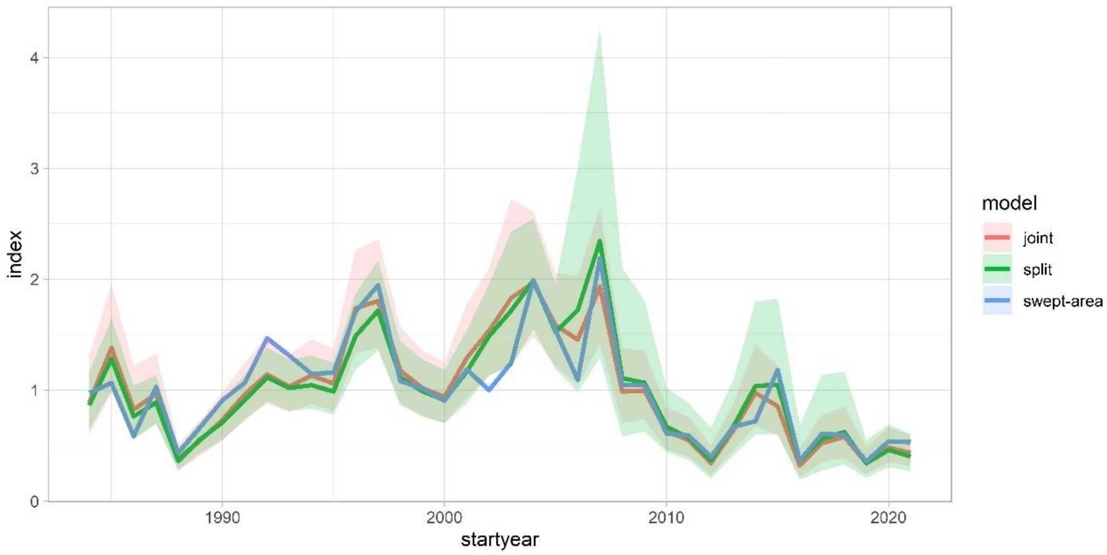


Figure 3.4.4. Total biomass indices, where blue is the swept-area index, green is the modelled index based on split time-series (1985-2005 and 2006-2021), while red is the modelled index based on the joint time-series 1984-2021.

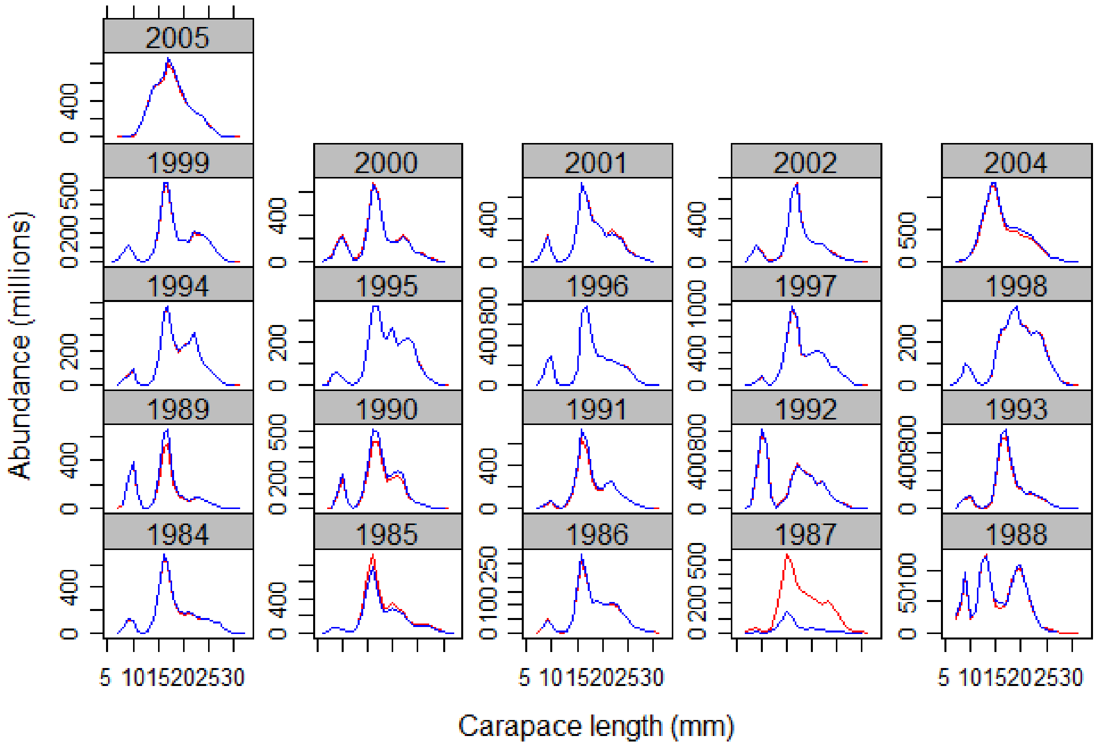


Figure 3.4.5. Abundance-at-length by year estimated through design-based approaches (StoX index: blue; SAS-index: red), for the time-series 1984-2002 and 2004-2005.

3.4.4 Biological information

Length Frequencies

Sweden

In previous assessments length frequencies were requested from landings and discards combined (back to 2008). In this benchmark length frequencies were requested from landings and discards separately. Therefore, lengths for discards and landings were recalculated for 2016-2020 as described in WD05 (Annex 2): Sampling design and estimation of commercial catches of northern shrimp *Pandalus borealis* by the Swedish bottom otter trawl fisheries in the 2016-2020.

Lengths were provided per year, quarter, and the area Skagerrak/Kattegat for 1990-2015 with methodology described in WD06 (Annex 2): Description of methodology behind compilation of Swedish data on *Pandalus borealis* (1908-2020) used in the *Pandalus* benchmark (WKPRAWN 2022).

Onboard length samples are collected from: unsorted catch, boiled shrimp, raw shrimp and discard. From self-sampling length samples are taken from unsorted catch. For the period 2005-2007 it is not known if samples were from self-sampling and/or from onboard sampling. For 1990-2004 samples were ordered from the fish auction in Smögen.

Denmark

The Danish length-data cover the period from 2013 to 2020, where a total of 62 trips and 172 hauls were sampled for subsequent length measurements in the laboratories of DTU Aqua (WD12). The length-data are based on three separate sample-fractions per haul; 1) small, discarded shrimp, 2) medium-sized shrimp landed raw, and 3) large shrimp boiled on-board before being landed. Each fraction typically has a sample-size of around 160-180 shrimp. The shrimp are length-measured individually (carapace length with a precision of 1 mm) and subsequently weighed together by each 1 mm group. The weights of the 1 mm length-groups are then summed to provide total sample weight by fraction.

The observer program provides on average about 4-5 samples of shrimp-lengths per quarter. The individual Length Frequency Distributions (LFDs) and weights of these samples are summed by each of the three catch-fractions before calculating a quarterly LFD per catch weight by fraction (the pooled LFD in absolute numbers to the pooled weight). These quarterly LFD/weight factors are raised to the total official catches within each of the three fractions (discarded, medium-sized, and large shrimp). Finally, these are summed to provide total quarterly DK catches in 3a and 4a-east in absolute numbers per 1 mm length group (carapace length) for the years from 2013-2020.

As the Danish observer-based sampling data are also almost exclusively from 3.a (only 5 of 65 trips are from 4.a East), the LFD data have been pooled for the two areas and the same quarterly LFDs are raised to the total catches from both areas.

Norway

Unsorted catches

Since 2005, selected Norwegian shrimp fishers have self-sampled their catches, providing shrimp samples enabling estimation of Norwegian catch at length (WD02). The samples (1-2 kg) are taken from the unsorted catch. The fishers are asked to take a sample once every month, but for various reasons not all follow this scheme. The samples are sent to IMR, where the shrimp are sorted by sex and maturity stage and length measured (CL). Weight of the catches from which samples are taken have only been provided since 2020. Since 2005, the Norwegian Coast Guard has, as part of their inspections of fisheries, collected shrimp samples from the fishery in the

Norwegian waters of Skagerrak and the Norwegian Deep, also from the unsorted catch. Numbers per length (whole mm length groups) are summed for all samples per year, quarter, and area. Numbers per length in the pooled samples are raised to numbers per length in total catches, i.e. corrected landings + estimated discards (per year, quarter and area), by assuming that the relation between the numbers per length in the samples and in the total catch is the same as the relation between the summed weight of the samples and the weight of the total catch (per area and quarter). Data from unsorted catches from 2006 to 2018 are used as input data to the benchmark.

The number of fishermen self-sampling their catches has varied between two and eight. The annual number of samples from both fishermen and the Coast Guard has also varied; in 2015-2020, between 77 and 135. In 3.a, 50-90% of the samples come from vessels <15 m, while in 4.a, 40-75% of the samples come from this vessel group. Within years, there have been some variations between quarters and areas, where 3.a is the best sampled area. In some years, coastal areas have been better covered by sampling than off-shore areas.

Sorted catches

The benchmark requested length data from discards and landings separately. From the Norwegian fishery, such data exist from 3.a since 2017, and from both 3.a and 4.a since 2019. These data come from the Norwegian Coastal Reference fleet (CRF) (see description of sampling program above under Norwegian discards). Norwegian length data from sorted catches were provided for both areas for 2019-2020. In these two years, 4.a was well covered by sampling from the CRF (two vessels), while 3.a was only poorly covered (Working document).

3.4.5 Length-weight and growth

Growth parameters were estimated internally by the model and were considered to be the same for both females and males.

A length–weight relationship (a = 0.0016, b = 2.7532) was estimated for the benchmark in 2016, from the Swedish commercial catch samples, where a subsample of shrimps was weighted from the total length-measured shrimps from trips for the years 2013 and 2014 (ICES 2016). The weight–length relationship was constant for all years in the stock assessment.

3.4.6 Maturity and size at sex change

Initially, the size at sex change, i.e. the size at which 50 percent of the individuals are females (L_{50}), was assumed to be equal to the size of maturity (of females), and the working document was written based on this assumption. L_{50} in the text below therefore refers to size at maturity (of females) as well as size at sex change. However, in the end it turned out that the results can only be used for size at maturity (see below).

Extensive analyses were conducted prior to the data evaluation workshop. These analyses used a range of different sources including Swedish commercial data and Norwegian survey data to not only investigate the annual variability in $_{50}$50, but also to identify the most appropriate available data to use for computing L_{50}. These analyses and results are detailed in *WD09_ Estimating size at sex change northern shrimp (Pandalus borealis)* and were presented at the data meeting.

Following the presentation, it was decided that the Norwegian survey data were the most appropriate data to use for the L_{50} estimates, as the survey also targets younger, smaller individuals. Since the survey has been conducted at three different times of the year since it started in 1984, estimates of L_{50} are not comparable throughout time. It was therefore also decided that estimates of L_{50}, and the associated slope of the logistic curve K, would only be made for the time period 2006-2021, during which the surey has been continiously conducted in January/Feburary. It was also decided that the analyses should be conducted for the North Sea and Skagerrak areas separately. If the analysis detected no clear trend in L_{50} or K between the two areas, it was decided that an average L_{50} and K for 2006 to 2021 would be used. It was also decided at the data meeting that an average of L_{50} and K across the two areas should also be presesented as an alternative to one of the models in the ensamble. For details of the analyses and results conducted prior to the data meeting see the *Annex* to the *WD09_ Estimating size at sex change northern shrimp (Pandalus borealis).* The analyses conducted after the data meeting and the final decisions made are presented below.

Data collected during the Norwegian survey were used to calculate yearly estimates of L_{50} and K. L_{50} was estimated by fitting a generalized logistic mixed model as described in WD 09. Figure 3.4.6 presents yearly estimates of L_{50} and K for the North Sea and Skagerrak. Although L_{50} and K were variable between years, there were no clear trends detected over time for either area, i.e. the regression lines were not significantly different from 0. Averages of L_{50} and K were therefore computed for each of these areas and, in turn, for the two areas combined (Table 3.4.3).

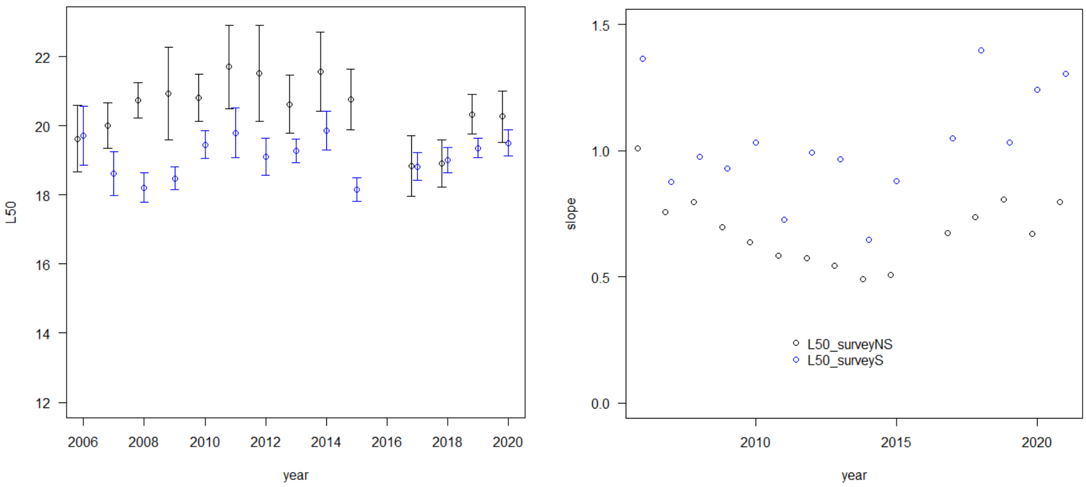


Figure 3.4.6. Yearly estimates of L_{50} (left panel) and K (right panel) for *Pandalus borealis* based on Norwegian survey data from 2006 to 2021 for the North Sea (NS in black) and Skagerrak (S in blue).

Table 3.4.3. Mean estimates of L_{50} (left panel) and K (right panel) for *Pandalus borealis* across the years 2006 to 2021 for the North Sea and Skagerrak and the mean across the two areas.

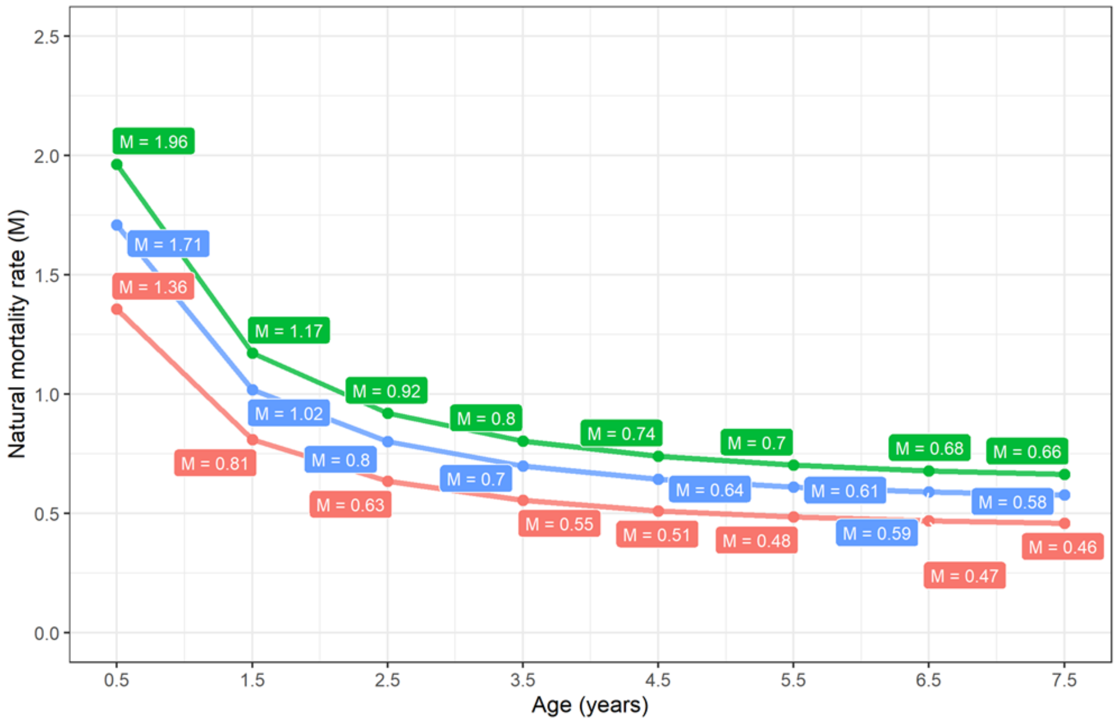
Area	L50	K
North Sea	20.38868	0.684876
Skagerrak	19.09369	1.027649
Mean of North Sea and Skagerrak	19.74119	0.8562624

For the resulting base case model, the benchmark group agreed to move forward with the assumption that the biology of the two areas is the same. Thus, the mean L_{50} and K from the two areas (Table 3.4.3) was used and was assumed to equal the size at maturity. As it turned out, the hermaphroditic parameters of the Stock Synthesis assessment model are defined in ages and not in lengths, consequently size at sex change was fitted by the model using the sex separated data (see below in the model specifications).

3.4.7 Natural mortality

Age-varying natural mortality rates (M) were estimated for both sexes using the methods described in the working document (WD04). In summary, the Barefoot Ecologist's Toolbox (http://barefootecologist.com.au/shiny_m) was used to derive a distribution of M values for *Pandalus* based on species-specific life-history traits. Independent of the methods used for estimation, M remains a highly uncertain parameter and to account for this, the distribution of M was summarised using the median, lower and upper quartiles (5% and 80% CIs respectively). An upper 80% CI was used as the upper quartile, as opposed to a 95% CI, to circumvent the long tail of the M distribution. These summaries were then translated into three age-based vectors using established methods, such that each age gets an associated M value (see WD07). The three M by age vectors are presented in Figure 3.4.7 and will be used as plausible scenarios for M in the final ensemble model (see section 3.4.11 for further details).

Figure 3.4.7. Northern shrimp in divisions 3.a and 4.a East. Age-specific natural mortality rates (M) used in the reference model. Shown in red, blue and green are the lower, median and upper scenarios, respectively.



Moreover, to reduce the number of parameters used in the model, M was set using 4 age breaks (not 8): age 0.5, 1.5, 3.5 and 7.5, where M for the adjacent ages is linearly interpolated using the values estimated for the age breaks (Table 3.4.4.).

Table 3.4.4. Northern shrimp in divisions 3.a and 4.a East. Natural mortality rates (M) by age breaks used in all three plausible M scenarios.

	Age			
Scenario	0.5	1.5	3.5	7.5
Median	1.709	1.019	0.699	0.577
Lower	1.356	0.809	0.554	0.458
Upper	1.964	1.171	0.803	0.664

In addition to the three scenarios detailed above, a fourth scenario for M was made prior to the meeting using predator abundances. The work behind this fourth scenario is detailed in WD10 and summarised in section 3.3. However, given the need for further research and development of the methodology, the meeting rejected the use of this index (both the 4 species and 9 species versions) in the assessment, i.e. it was not included in the ensemble model as a plausible scenario for M.

3.4.8 Assessment model

In the following text, the base case and final assessment model for *Pandalus* in the Skagerrak and Norwegian Deep area are described. The new models advance the current assessment (described in Benchmark report 2016) by incorporating the species' hermaphroditic life-history, as well as the inclusion of catches since 1908, modelling of discards through time, the separation of the stock in two areas (3.a and 4.a East, respectively), and the separation of the fishery into six fleets (by nation and area). Further, in section 3.4.11 a new ensemble model is described. The ensemble is by far the biggest novelty used in this benchmark assessment, and its inclusion allows multiple plausible models (and parameter sets) to be tested within a single integrated framework. The ensemble approach provides a more robust quantification of model uncertainty and more reliable predictions of stock status.

Assessment model for *Pandalus*

Assessment of Northern shrimp (*Pandalus borealis*) in divisions 3.a and 4.a East is conducted using the Stock Synthesis (SS) model (Methot & Wetzel, 2013; Methot *et al*., 2021). Stock Synthesis is programmed in the ADMB C++ software and searches for the set of parameter values that maximize the goodness-of-fit, then calculates the variance of these parameters using inverse Hessian and MCMC methods. The assessment was conducted using the 3.30.18 version of the Stock Synthesis software under the windows platform.

The total likelihood of the model is composed of several components, including the fit to survey and CPUE indices, tag recovery data (when tagging data are used), fishery and survey length frequency distributions (LFDs), age compositions (when present), conditional age-at-length compositions and catch data. There are also contributions to the total likelihood from the recruitment deviates and priors on the individual model parameters (if any). The model is configured to fit the catch so the catch component of the likelihood is generally small (although catch penalties might be created). Details of the formulation of the individual components of the likelihood are provided in Methot & Wetzel (2013).

Data

The following landings and discard data were provided by country and were inputted into in the assessment. All input data was spilt by quarter and area:

Swedish landings

- 1978-2020 landings data were taken from the NIPAG report (ICES 2021). These were split by quarter and area using estimated yearly proportions from Swedish national sources. Landings were also corrected for boiling based on the approach detailed in WD05.
- 1963-1977 landings data were taken from the ICES historical database (https://www.ices.dk/data/dataset-collections/Pages/Fish-catch-and-stock-assessment.aspx). These were split by quarter using average proportions estimated from Swedish national sources 1978-1982 (5-year average), and split by area based on proportions taken from the ICES historical database. Landings were also corrected for boiling.
- 1908-1962 landings data were taken from the ICES historical database. These were split by quarter using average proportions estimated from Swedish national sources 1978-1982 (5-years average), and split by area based on proportions taken from the ICES historical database. Landings were not corrected for boiling.

Swedish discards

- 2008-2015 discards data were taken from the ICES summary sheet. These were split by quarter and area using estimated yearly proportions as derived from Swedish landings.
- 2016-2020 discards data were calculated by area and quarter. The method for this calculation is described in working document: Sampling and design-based estimation of commercial catches by the Swedish bottom otter trawl fisheries operating in ICES subdivision 27.3.a.21 in the 2016-2020.

Danish landings

- 2001-2020 landings data split by quarter and area were taken from Danish national sources. Landings were corrected for boiling.
- 1988-2000 landings data split by quarter and area were taken from Danish national sources. Landings were not corrected for boiling.
- 1908-1987 landings data were taken from the ICES historical database. These were split by quarter using average proportions as estimated from Danish national sources from 1988-1992 (5-years average), and split by area based on proportions taken from the ICES historical database. Landings were not corrected for boiling.

Danish discards

- 2013-2020 discards data split by quarter and area were taken from Danish national sources.
- 2009-2012 discards data were taken from the ICES summary sheet. These were split by quarter and area using the average proportions as estimated from Danish national sources 2013-2015 (3-years average).

Norwegian landings and discards

- 1908-2020 landings and discards data split by quarter and area were taken from Norwegian national sources.

All data inputs to the assessment model are listed in Table 3.4.5.

Table 3.4.5. Northern shrimp in divisions 3.a and 4.a East. Input data used in the Stock Synthesis assessment model.

Type	Name	Year range	Range
Landings	Total annual landings in tonnes	1908-2020	
Discards	Total annual discards in tonnes	Danish fleet 2013-2020 Norwegian fleet 2009-2020 Swedish fleet 2008-2020	
Length frequency distributions (LFDs)	Catch in numbers per length class and sex	Danish fleet 2013-2020 Norwegian fleet 2006-2020 Swedish fleet 1990-2020 Survey: 1984-2020 (excluding 2003)	0.2–3.5 cm
Maturity ogives	Empirical maturity at length estimated from survey data		
Natural mortality	Natural mortality by age class assumed to be costant for the entire time-series		0-8+
Surveys indices	Biomass index from survey by area	Survey: 1984-2020	
SSB index	SSB proportional to fecundity		

Sample sizes, CVs and data weighting

For the commercial fleets, the coefficient of variation (CV) of the catches (including the discards) was set to 0.2. The CV of the initial catches of the commercial fleets was also set to 0.2. The annual sample size associated with the LFD data is reported as the number of trips sampled for commercial catches (as reported from national sources) and the number of hauls for the surveys. Sample size (i.e. number of trips) by quarter for the years 2005 to 2015 for Swedish LFDs in 3a is estimated as the average sample size by quarter as observed for years 1990-2004. The CV of the surveys was available and had an average of 0.2 over the entire time-series. No weighting of the LFDs was used in the base case model.

3.4.9 Model settings

Base model

The assessment model of *Pandalus* in divisions 3.a and 4.a East is a two area, length-based model with a population comprised of 8+ age classes (with age 8 representing a plus group). The population is spilt into two sexes, with hermaphroditic individuals being born as males and changing to female later in life (protandrous hermaphroditism). The model has a quarterly time step to account for differences in individual growth throughout the year.

The model starts in 1970 (although the same model configuration starting in 1908 was also tested) and the age structure of the initial population was assumed to be in an exploited state. The initial catches were assumed to be the average of the preceding five years (1965-1969). Fishing mortality was modelled using a fleet-specific method (Methot *et al.*, 2021). Option 5 was selected for the F report basis; this option corresponds to the fishing mortality requested by the ICES framework (i.e. simple unweighted average of the F of the age classes chosen to represent the F_{bar} (age 1-3)).

Spawning stock biomass was estimated at the beginning of the year and it was considered proportional to fecundity. In the model, recruitment was assumed to be a single event occurring at the beginning of the year. Recruitment was derived from a Beverton and Holt (BH) stock recruitment relationship (SRR) and variation in recruitment was estimated as deviations from the SRR. Recruitment deviates were estimated for 1984 to 2020 as main recruitment deviations (37 annual deviations) and for 1978 to 1983 as early recruitment deviation (6 annual deviations). Recruitment deviates were assumed to have a standard deviation (σR). σR is the stochastic recruitment process error and the estimation of this parameter within integrated models is generally recognised to be problematic (Kolody *et al.*, 2019). Consequently, σR for individual recruitment estimates were fixed at values large enough to prevent the SSR from constraining individual recruitment estimates (e.g. analogous to traditional VPA) (Kolody *et al.*, 2019). A meta-analysis of the estimation of σR was performed outside the operative model (ISSF, 2011) and yielded a median estimate between 0.2 and 0.5. This suggests that σR is often inflated in assessment models. In the base case σR is estimated within the model. The steepness (h) for the SRR and the autocorrelation of recruitment is also estimated within the model.

Growth parameters were estimated internally by the model and were considered to be the same for both females and males. Weight was estimated from a length-weight relationship (a = 0.0016, b = 2.7532) while length-at-maturity was described by a sigmoidal function with $L_{50\%}$ set at 1.974 cm. Length-weight and length-at-maturity parameters were fixed and derived externally using commercial and survey data respectively (sections 3.4.5 and 3.4.6.). Further details on how weight and length at maturity were derived are included in the stock annex and in the WD09, respectively.

Age-varying M's for both sexes were estimated based on the methods described in working document WD04, a summary of this estimation procedure is provided in section 3.4.2.4. In total, three age-varying M scenarios were proposed as plausible and are integrated within the final ensemble model (see section 3.4.2.5.2).

Exploratory runs

A series of exploratory runs were tested prior to the meeting with the aim of exploring what effect different model configurations might have on model behaviour and resulting outputs. All exploratory runs were based on the base model described above. The following configurations were explored and were presented to the meeting (Table 3.4.6):

Table 3.4.6. Northern shrimp in divisions 3.a and 4.a East. Alternative configurations of the assessment model.

Name	Brief description	Reason
HIST	Model starts from 1908	Including historical catches in the model (1908-1969)
Areasexratio	Sex ratio used to split LFDs into males and females is area specific	Include area specific sex ratio
Hermfix	Hermaphroditic curve fixed using expert knowledge	Exploring different shapes of hermaphroditic curve
TV_D_ogives	Discard retentions are time varying for all fleets (except NOR4a) from 2017	Exploring different discard patterns
Reference_no2016	Removing 2016 from the survey time-series	Spatiotemporal coverage of the survey in 2016 differs from the other years

The addition of historical catches (*HIST*) did not change the fit of the model and led to an equivalent model to the base case (Table 3.4.7, Figure 3.4.9a). *Areasexratio* and *Hermfix* configurations did differ from the base case (bigger relative change in Figure 3.4.8), however, both led to a reduction in fit to LFDs and surveys (Table 3.4.7). The Hermfix configuration was the only run that produced an F trend that differed from all other exploratory runs (Figure 3.4.9b). The TV_D_ogives configuration did improve the fit of the model but the impact on the results was minimal (Table 3.4.7, Figure 3.4.9b). The Reference_no2016 configuration also had no considerable improvement on model fit nor an impact on the results (Table 3.4.7, Figure 3.4.9b).

Table 3.4.7. Northern shrimp in divisions 3.a and 4.a East. Summary table of exploratory runs and their estimated parameters

Label	Type	Reference	HIST	Areasexratio	Hermfix	TV_D_ogives	Reference_no2016
Total_likelihood		1559.4	1559.4	1716.2	1765.7	1442.4	1545.9
Survey_likelihood							
Catch_likelihood		177.0	177.0	178.9	184.5	116.1	171.7
Discard_likelihood		480.9	480.9	486.7	484.3	452.6	478.1
	ALL	-56.7	-56.7	-54.8	-49.8	-64.9	-70.1
	NORSURVEY3a	-30.7	-30.7	-29.3	-26.9	-36.9	-37.2
	NORSURVEY4a	-26.0	-26.0	-25.5	-22.9	-28.1	-32.9
Length_likelihood							
	ALL	962.4	962.4	1114.2	1151.9	945.4	970.0
	DK3a	107.5	107.5	110.9	119.9	109.5	108.4
	DK4a	81.0	81.0	80.0	89.1	87.1	81.0
	NOR3a	86.7	86.7	88.6	97.8	78.0	87.9
	NOR4a	50.0	50.0	50.1	47.2	49.2	49.2
	SWE3a	135.5	135.5	134.1	145.0	134.6	136.6
	SWE4a	14.6	14.6	14.6	15.3	15.1	14.7
	NORSURVEY3a	246.8	246.8	335.8	343.5	238.6	249.2
Parameters							
	NatM_break_1_Fem_GP_1	1.709	1.709	1.709	1.709	1.709	1.709
	NatM_break_2_Fem_GP_1	1.019	1.019	1.019	1.019	1.019	1.019
	NatM_break_3_Fem_GP_1	0.699	0.699	0.699	0.699	0.699	0.699
	NatM_break_4_Fem_GP_1	0.577	0.577	0.577	0.577	0.577	0.577
	L_at_Amin_Fem_GP_1	0.7068	0.7068	0.7106	0.0000	0.7030	0.7067
	L_at_Amax_Fem_GP_1	2.6941	2.6942	2.7310	2.7926	2.6893	2.6933
	VonBert_K_Fem_GP_1	0.455	0.455	0.439	0.348	0.458	0.455
	CV_young_Fem_GP_1	0.119	0.119	0.120	0.121	0.119	0.119
	CV_old_Fem_GP_1	0.044	0.044	0.040	0.078	0.044	0.045
	Wtlen_1_Fem_GP_1	0.002	0.002	0.002	0.002	0.002	0.002
	Wtlen_2_Fem_GP_1	2.753	2.753	2.753	2.753	2.753	2.753
	Mat50%_Fem_GP_1	1.974	1.974	1.974	1.974	1.974	1.974
	Mat_slope_Fem_GP_1	-10.253	-10.253	-10.253	-10.253	-10.253	-10.253
	Wtlen_1_Mal_GP_1	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016
	Wtlen_2_Mal_GP_1	2.7532	2.7532	2.7532	2.7532	2.7532	2.7532
	Herm_Infl_age	2.69E-06	2.69E-06	0.48987	2.6	2.41E-06	2.42E-06
	Herm_stdev	3.351	3.351	2.902	0.400	3.381	3.348
	Herm_asymptote	1	1	1	1	1	1
	RecrDist_GP_1_area_2_month_1	0.252	0.251	0.125	-0.028	0.227	0.225
	RecrDist_GP_1_area_2_month_1_dev_se	1.5	1.5	1.5	1.5	1.5	1.5
	RecrDist_GP_1_area_2_month_1_dev_autocorr	0	0	0	0	0	0
	SR_LN(R0)	17.952	17.953	17.831	18.726	17.954	17.892
	SR_BH_steep	0.70	0.69	0.77	0.76	0.71	0.73
	SR_sigmaR	0.50	0.50	0.44	0.49	0.43	0.51
	SR_autocorr	0.32	0.32	0.35	0.39	0.26	0.33
Derived quantities							
	SB0	43919	43973	39353	40494	43900	41323
	SSB_2020	5702	5701	5700	6648	5842	5775
	Bratio_2020	0.37	0.37	0.41	0.47	0.33	0.40
	MSY_SPR	11825	11953	10910	10114	11985	11422
	F_SPR	0.23	0.23	0.23	0.09	0.23	0.23

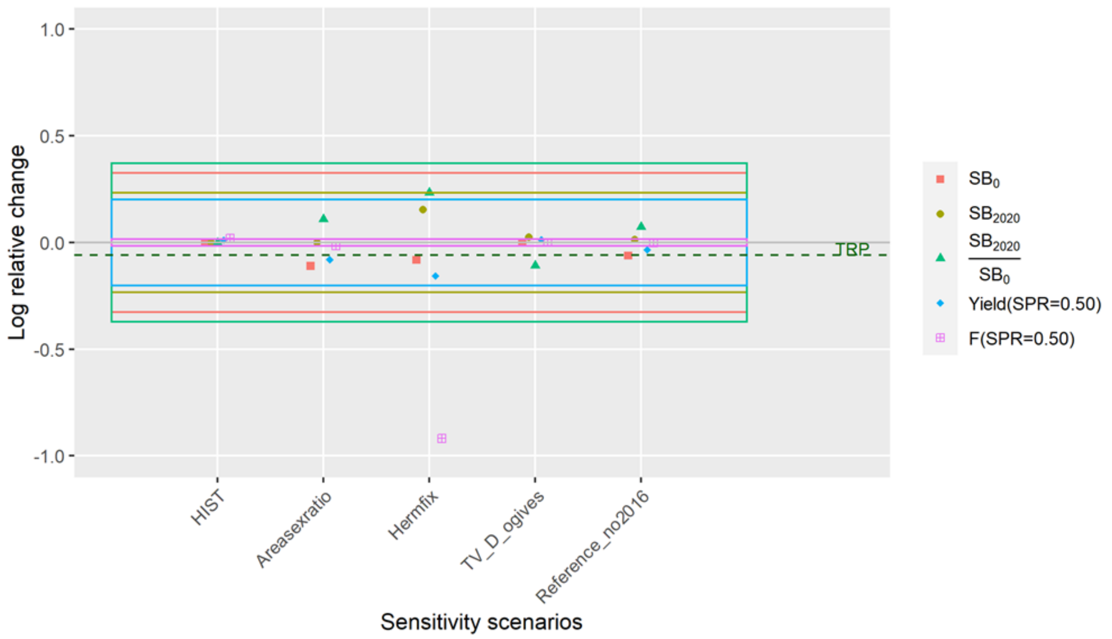


Figure 3.4.8. Northern shrimp in divisions 3.a and 4.a East. Sensitivity plot of exploratory runs. The Log relative change refers to changes between each alternative configuration and the base case: 0 means no change. Changes are calculated on a set of predefined yield, biomass, and F model outputs. Coloured horizontal lines represent 95% confidence intervals.

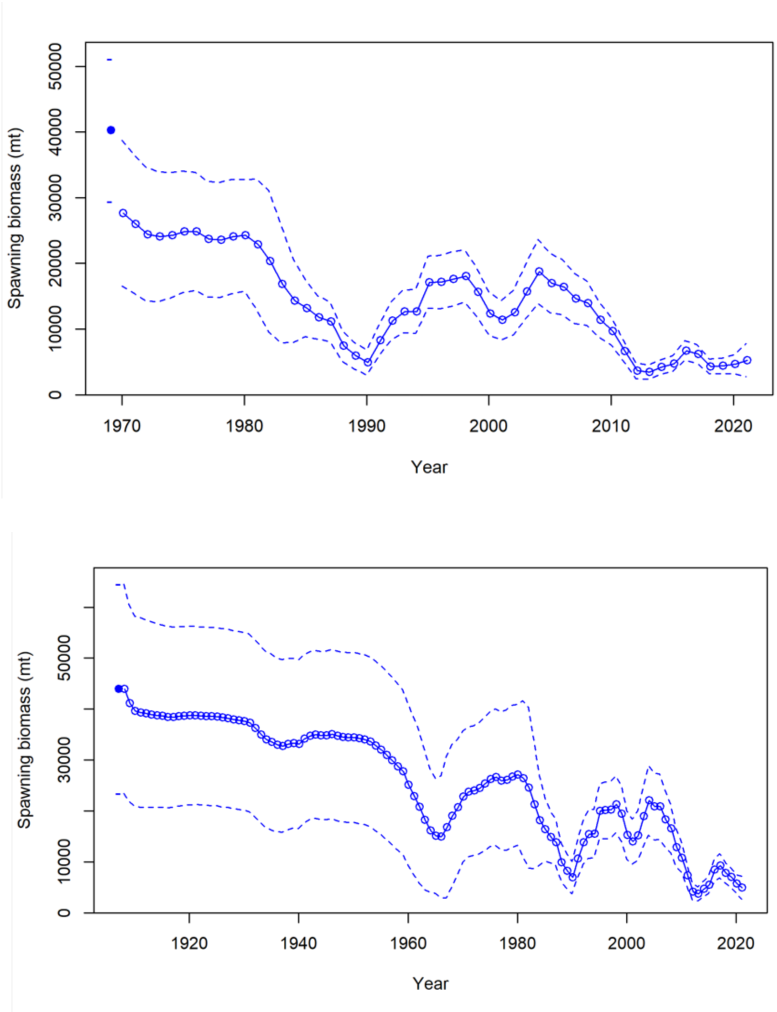


Figure 3.4.9a. Northern shrimp in divisions 3.a and 4.a East. Summary plots of the exploratory runs. Comparing SSB of the reference run (top) with SSB from the HIST exploratory run (see Table 3.4.5 for an explanation of the exploratory runs)

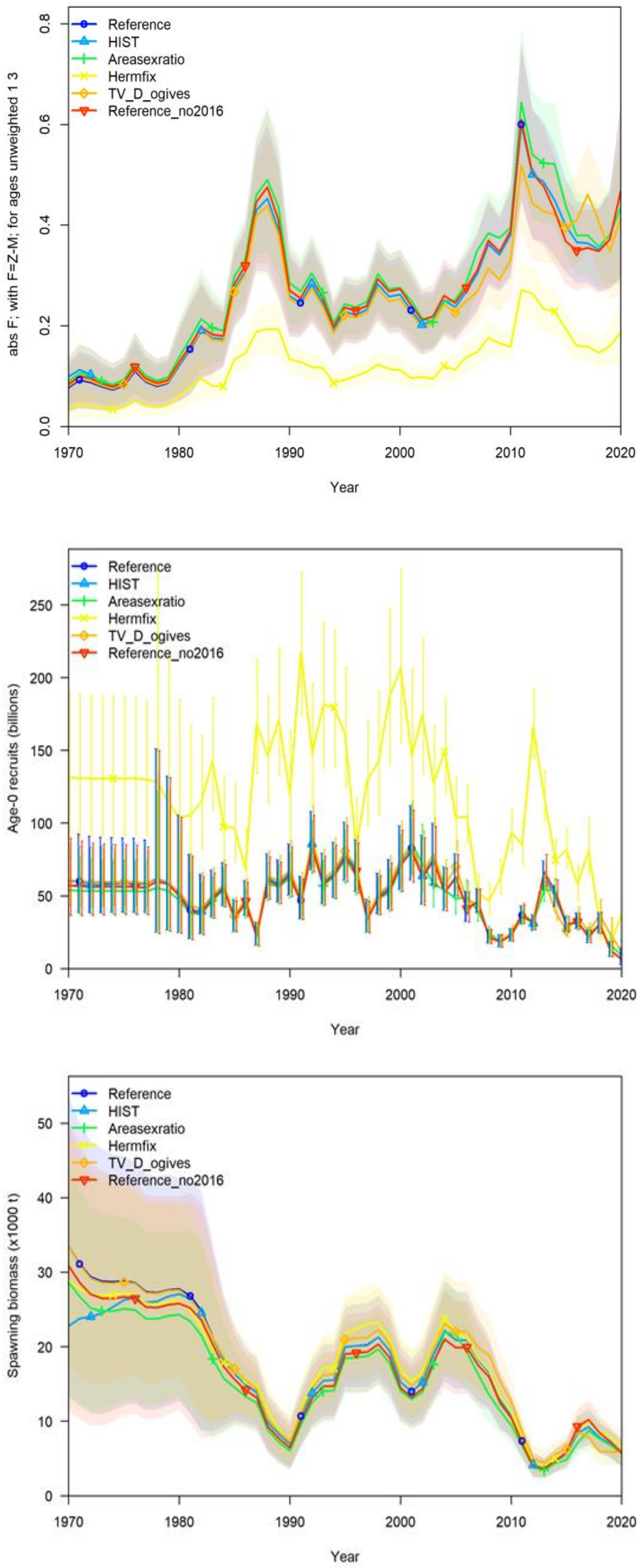


Figure 3.4.9b. Northern shrimp in divisions 3.a and 4.a East. Summary plots of the exploratory runs (see Table 3.4.5 for an explanation of the exploratory runs).

3.4.10 The final model

After careful scrutiny of the different model configurations tested (see section above), the meeting decided to select the model with time varying discard ogives (*TV_D_ogives*) as the final model. This decision was mainly due to an overall better fit of the model to the data (Table 3.4.7). It was also decided the 2016 survey would be retained in the model (see section 3.4.3 for further discussion on this). The configuration of the final model is reported in Table 3.4.8.

All parameter estimates and variances were reasonably well estimated by the final model (i.e. CV < 1) except for *Herm_Infl_age* and some initial *F*s which were estimated at the lower bound. An overview of the data included in the final model is shown in Figure 3.4.10. Model diagnostics for the final model are presented in figures 3.4.11-17. Jittering and MCMC are shown in Figures 3.4.18-19, however, these consider the base case model and not the final model due to long run times. Results from retrospective analysis of the final model, as well as hindcasting are shown in figures 3.4.20-21 and Table 3.4.9. All diagnostics used to compare the different model configurations and test the final model were considered valid for this purpose as shown by the published work of Carvalho *et al.* (2021) and Kell *et al*. (2021). For a thorough description of the MCMC analysis, jittering procedure, retrospective analysis, and hindcasting we refer the reader to WD03.

Further, estimates of surplus production were used to check whether predictions about changes in biomass can be made reliably based on catch and current biomass (clockwise or linear behaviour) or whether there is evidence of non-stationarity in production processes, i.e. are the dynamics driven by climate and oceanic conditions (counter clockwise; Walters et al. 2008). In the case of *Pandalus* in divisions 3.a and 4.a East, Figure 3.4.22 shows a general clockwise pattern indicating that changes in stock biomass can be made reliably based on catch and current biomass. The production function (Figure 3.4.23) is typically left skewed and flat-topped, which implies that fishing at the naïve F_{MSY} brings the stock close to B_{lim} with very reduced theoretical gain in yields compared to more conservative F_{MSY} proxies such as $B_{30\%}$ or $B_{35\%}$.

Finally, when all diagnostic tests are considered together, the power to detect model misspecification improves without a substantial increase in the probability of incorrectly rejecting a correctly specified model (Carvalho *et al.*, 2017, 2021). Consequently, all the diagnostics used here should be all applied routinely. When the criterion for rejecting a model is a failure of at least one of the diagnostic tests, nearly 90% of most misspecifications are detected with no real increase in the probability of a false detection (Carvalho *et al.*, 2017, 2021). Residual analyses were easily the best detector of misspecification in the observation model, while the retrospective analysis had low rates of detection of mis-specified models (Carvalho et al., 2017, 2021), although retrospective analysis is effective in detecting un-modelled temporal variation (Hurtado-Ferro et al., 2014). Opposed to the widely used maximum-likelihood estimator, MCMC gives clear warning signs when a non-identifiable model is used for fitting (Siekmann *et al*., 2012). In this context, we created a table that summarises all diagnostics for the final model and compare it with the base case proposed before the benchmark (now called "Discards ogive not time varying"; Table 3.4.10). Table 3.4.10 is an attempt to sum up a multidimensional space and thus it needs to be seen as a guidance more than as a definitive result. However, it is evident that the final model performs well in most of the key diagnostic tests performed. Thus, the final model was accepted by the WKPRAWN meeting and proposed as the model to be used to integrate the key dimensions of uncertainty in the final ensemble (Section 3.4.11).

Table 3.4.8. Northern shrimp in divisions 3.a and 4.a East. Settings of the final SS3 model. The table columns (left to right) show: parameter name, number of estimated parameters, the initial values (from which the numerical optimization is started), the intervals allowed for the parameters, the priors used (value and standard deviation), the value estimated

by the model and its standard deviation. Parameters in bold are set and not estimated by the model. A * next to the parameter name indicates that the parameter is close to the bound.

Parameter	Number estimated	Initial value	Bounds (low, high)	Prior	Value (MLE)	Standard deviation
<u>Natural mortality (M)</u> (age classes 0.5, 1.5, 3.5, 7.5)		**1.709, 1.019, 0.699, 0.577**				
<u>Growth both sexes</u>						
L_at_Amin	1	0.7	(0, 4)	No_prior	0.7	0.01
L_at_Amax	1	2.67	(2, 4)	No_prior	2.69	0.04
VonBert_K_young	1	0.47	(0.2, 0.8)	No_prior	0.46	0.02
CV_young	1	0.12	(0.005, 0.4)	No_prior	0.12	0.004
CV_old	1	0.05	(0.005, 0.4)	No_prior	0.04	0.009
<u>Length-weight both sexes</u>						
Wtlen_1		**0.0016**				
Wtlen_2		**2.7532**				
<u>Maturity at length for females</u>						
Mat50%		**1.97**				
Mat_slope		**-10.3**				
Eggs/kg_inter		**1**				
Eggs/kg_slope_wt		**0**				
<u>Hermaphroditism</u>						
Herm_Infl_age*	1	1.08	(0, 8)	No_prior	0.000002	0.007
Herm_stdev	1	2.33	(0.1, 21.3)	No_prior	3.38	0.2
Herm_asymptote		**1**				
<u>Recruitment distribution</u>						
RecrDist_GP_1_area_1_month_1		**0**				
RecrDist_GP_1_area_2_month_1	1	0.23	(-35, 25)	No_prior	0.23	0.16
<u>Cohort growth dev base</u>						
CohortGrowDev		**1**				
<u>Fraction female</u>						

Parameter	**Number estimated**	**Initial value**	**Bounds (low, high)**	**Prior**	**Value (MLE)**	**Standard deviation**
FracFemale		**1e-06**				
<u>Recruitment distribution</u>						
RecrDist_GP_1_area_2_month_1_dev_se		**1.5**				
RecrDist_GP_1_area_2_month_1_dev_auto-corr		**0**				
<u>Stock and recruitment</u>						
Ln(R_0)	1	17.96	(3, 30)	No_prior	17.95	0.19
Steepness (h)	1	0.69	(0.1, 1)	No_prior	0.71	0.12
Recruitment variability (σ_R)	1	0.53	(0, 2)	No_prior	0.43	0.06
Ln (Main recruitment deviation): 1984 - 2020	37					
Ln (Early recruitment deviation): 1978 - 1983	6					
Recruitment autocorrelation	1	0.26	(0, 1)	No_prior	0.26	0.12
<u>Initial catches</u>		Average of 1965-1969				
<u>Initial fishing mortality</u>						
InitF_seas_1_flt_1DK3a*	1	0.04	(0, 8)	No_prior	0.04	0.015
InitF_seas_1_flt_3NOR3a*	1	0.07	(0, 8)	No_prior	0.07	0.024
InitF_seas_1_flt_4NOR4a	1	0.08	(0, 8)	No_prior	0.08	0.029
InitF_seas_1_flt_5SWE3a	1	0.12	(0, 8)	No_prior	0.12	0.041
InitF_seas_2_flt_1DK3a*	1	0.06	(0, 8)	No_prior	0.06	0.022
InitF_seas_2_flt_3NOR3a*	1	0.07	(0, 8)	No_prior	0.07	0.023
InitF_seas_2_flt_4NOR4a	1	0.08	(0, 8)	No_prior	0.08	0.029
InitF_seas_2_flt_5SWE3a	1	0.16	(0, 8)	No_prior	0.15	0.053
InitF_seas_3_flt_1DK3a*	1	0.07	(0, 8)	No_prior	0.07	0.022
InitF_seas_3_flt_3NOR3a*	1	0.05	(0, 8)	No_prior	0.05	0.018
InitF_seas_3_flt_4NOR4a*	1	0.04	(0, 8)	No_prior	0.04	0.014
InitF_seas_3_flt_5SWE3a	1	0.14	(0, 8)	No_prior	0.13	0.046
InitF_seas_4_flt_1DK3a*	1	0.04	(0, 8)	No_prior	0.04	0.014

Parameter	Number estimated	Initial value	Bounds (low, high)	Prior	Value (MLE)	Standard deviation
InitF_seas_4_flt_3NOR3a*	1	0.04	(0, 8)	No_prior	0.04	0.015
InitF_seas_4_flt_4NOR4a*	1	0.02	(0, 8)	No_prior	0.02	0.007
InitF_seas_4_flt_5SWE3a	1	0.09	(0, 8)	No_prior	0.09	0.031
Selectivity (logistic)						
Commercial fleets						
Size_inflection_DK3a(1)	1	1.6	(0.35, 3.45)	No_prior	1.59	0.03
Size_95%width_DK3a(1)	1	0.3	(0.1, 3.45)	No_prior	0.34	0.02
Retain_L_infl_DK3a(1)	1	1.1	(0.2, 30)	No_prior	1.13	0.07
Retain_L_width_DK3a(1)	1	0.2	(-5, 30)	No_prior	0.23	0.03
Size_inflection_DK4a(2)	1	1.5	(0.35, 3.45)	No_prior	1.49	0.03
Size_95%width_DK4a(2)	1	0.3	(0.1, 4.5)	No_prior	0.34	0.03
Retain_L_infl_DK4a(2)	1	1.2	(0.2, 30)	No_prior	1.31	0.06
Retain_L_width_DK4a(2)	1	0.2	(-5, 30)	No_prior	0.22	0.03
Size_inflection_NOR3a(3)	1	1.6	(0.35, 3.45)	No_prior	1.64	0.02
Size_95%width_NOR3a(3)	1	0.3	(0.1, 4.5)	No_prior	0.35	0.02
Retain_L_infl_NOR3a(3)	1	1.3	(0.2, 30)	No_prior	1.26	0.05
Retain_L_width_NOR3a(3)	1	0.1	(-5.1, 30)	No_prior	0.15	0.02
Size_inflection_NOR4a(4)	1	1.6	(0.35, 3.45)	No_prior	1.62	0.04
Size_95%width_ NOR4a(4)	1	0.4	(0.1, 4.5)	No_prior	0.37	0.04
Retain_L_infl_ NOR4a(4)	1	0.7	(0.2, 30)	No_prior	0.75	0.27
Retain_L_width_ NOR4a(4)	1	0.3	(-5.1, 30)	No_prior	0.25	0.07
Size_inflection_SWE3a(5)	1	1.8	(0.35, 3.45)	No_prior	1.76	0.03
Size_95%width_ SWE 3a(5)	1	0.5	(0.1, 4.5)	No_prior	0.36	0.02
Retain_L_infl_ SWE 3a(5)	1	1.7	(0.2, 30)	No_prior	1.67	0.02
Retain_L_width_ SWE 3a(5)	1	0.1	(-5.1, 30)	No_prior	0.15	0.02
Size_inflection_ SWE 4a(6)	1	1.5	(0.35, 3.45)	No_prior	1.50	0.07
Size_95%width_ SWE 4a(6)	1	0.4	(0.1, 4.5)	No_prior	0.34	0.07
Retain_L_infl_ SWE 4a(6)	1	1.6	(0.2, 30)	No_prior	1.69	0.05

Parameter	**Number estimated**	**Initial value**	**Bounds (low, high)**	**Prior**	**Value (MLE)**	**Standard deviation**
Retain_L_width_ SWE 4a(6)	1	0.1	(-5.1, 30)	No_prior	0.11	0.03
Surveys						
Size_inflection_NORSURVEY3a(7)	1	1.6	(0.35, 3.45)	No_prior	1.6	0.04
Size_95%width_NORSURVEY3a(7)	1	0.5	(-5, 4.5)	No_prior	0.5	0.02
Size_inflection_NORSURVEY4a(8)	1	1.6	(0.35, 3.45)	No_prior	1.6	0.03
Size_95%width_NORSURVEY4a(8)	1	0.4	(0.1, 4.5)	No_prior	0.4	0.02
Catchability						
Survey 3a (floating option)						
Ln(Q) – catchability		**-0.594**				
Survey 3a (floating option)						
Ln(Q) – catchability		**-0.681**				

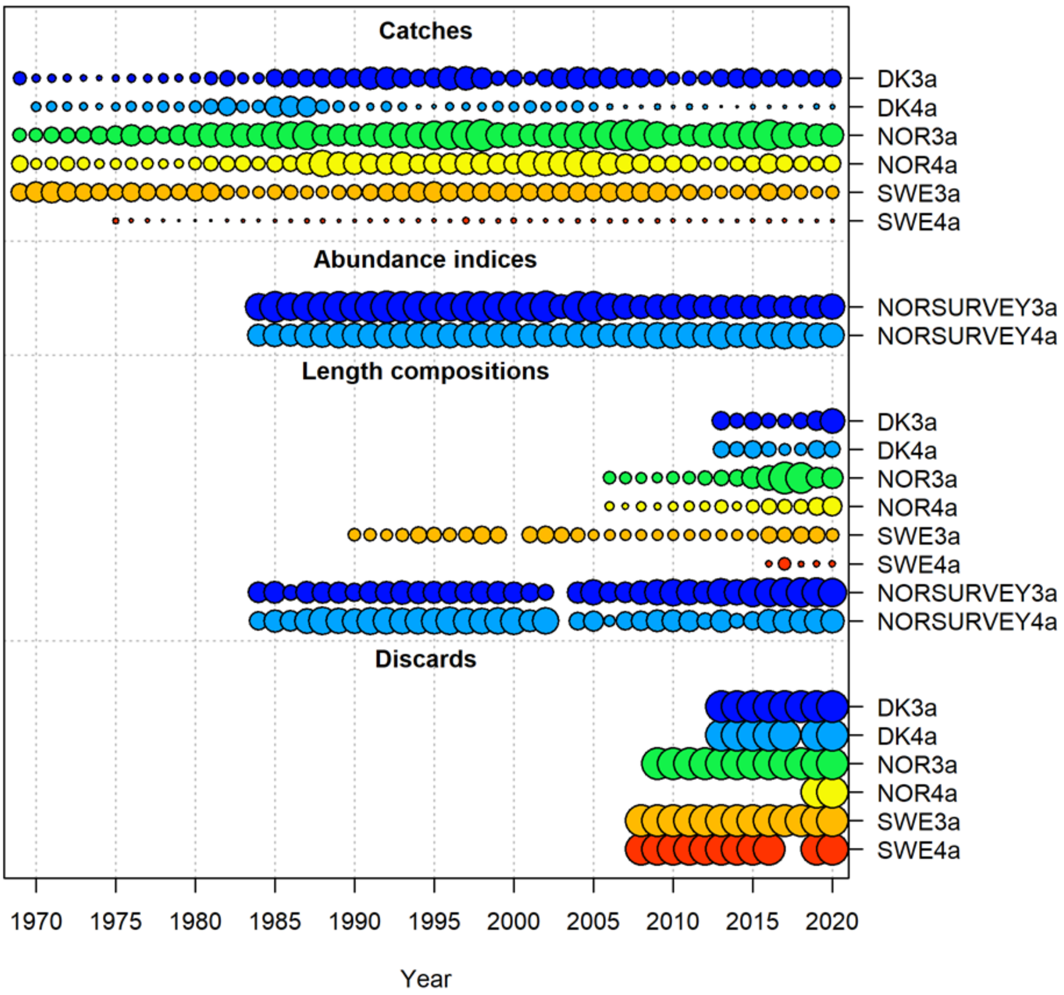


Figure 3.4.10. Northern shrimp in divisions 3.a and 4.a East. Summary of the input time-series included in the final model.

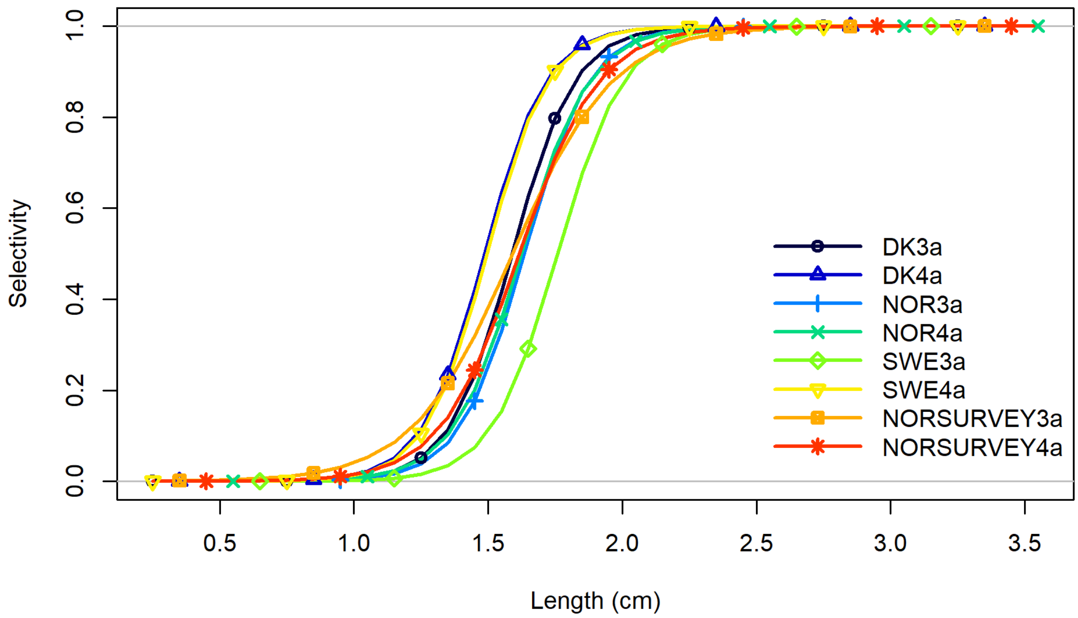


Figure 3.4.11. Northern shrimp in divisions 3.a and 4.a East. Length based selectivity by fleet.

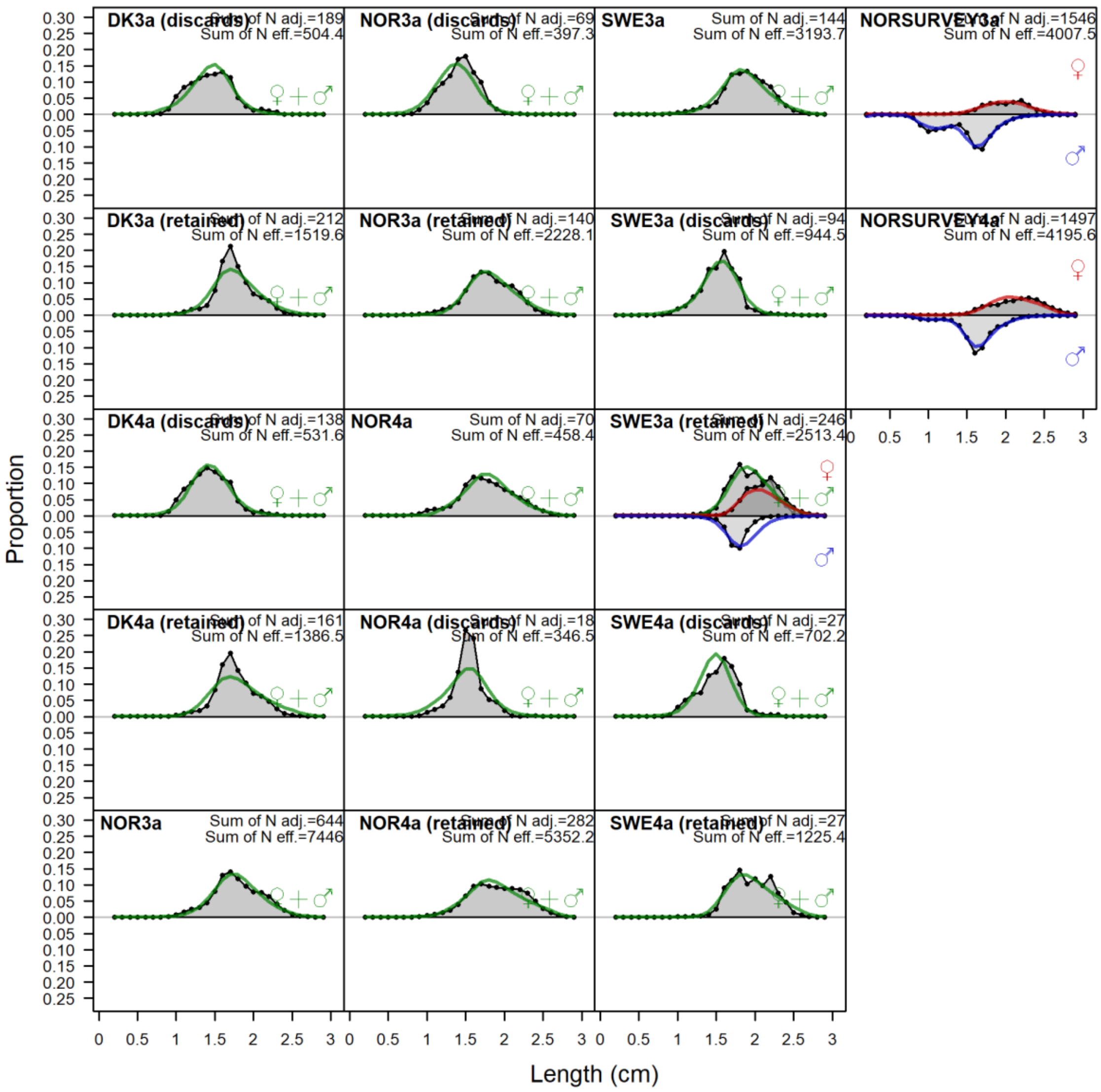


Figure 3.4.12. Northern shrimp in divisions 3.a and 4.a East. Model fits to LFD data.

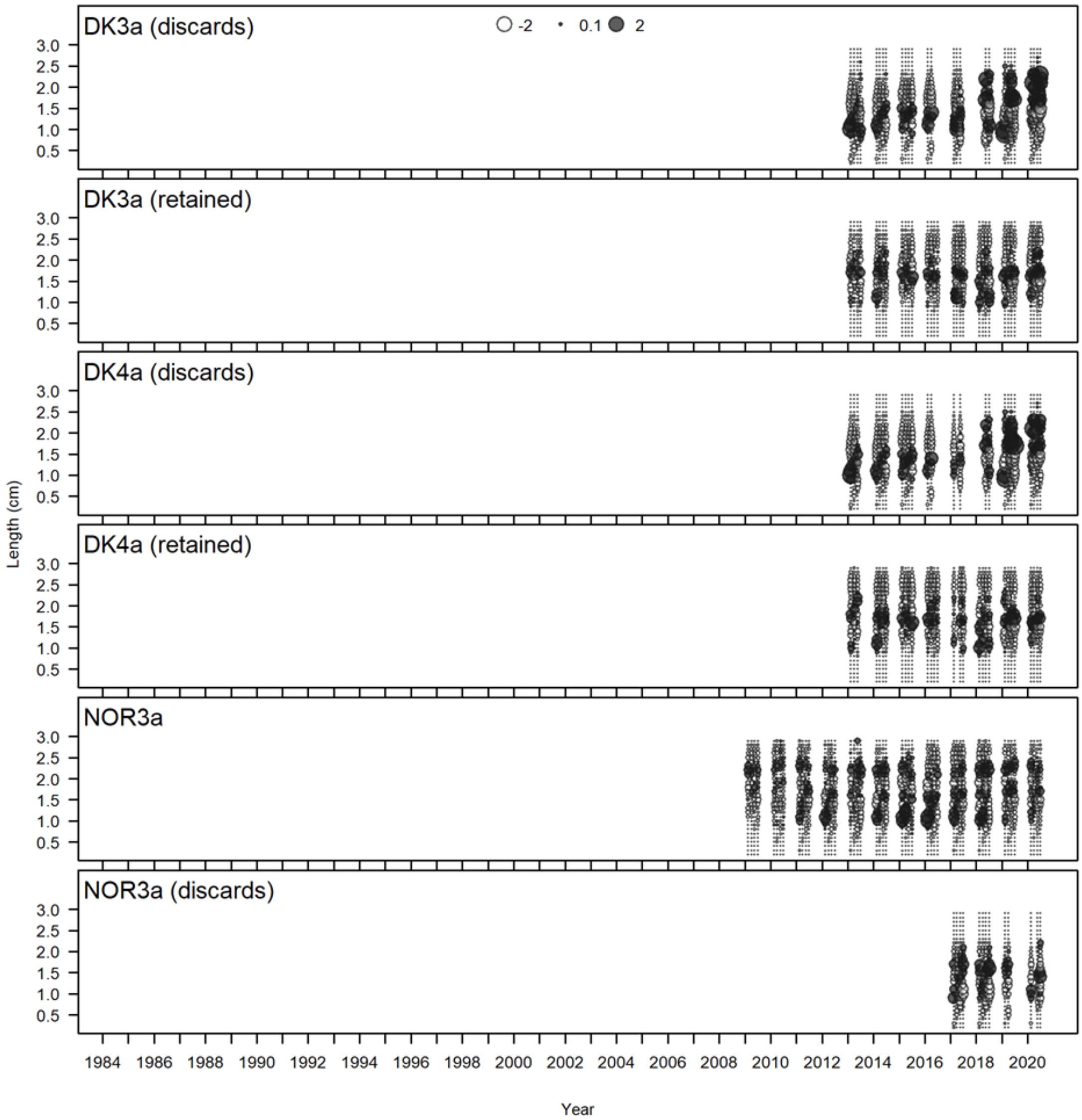


Figure 3.4.13. Northern shrimp in divisions 3.a and 4.a East. Residuals of fits to LFD data for the different fleets.

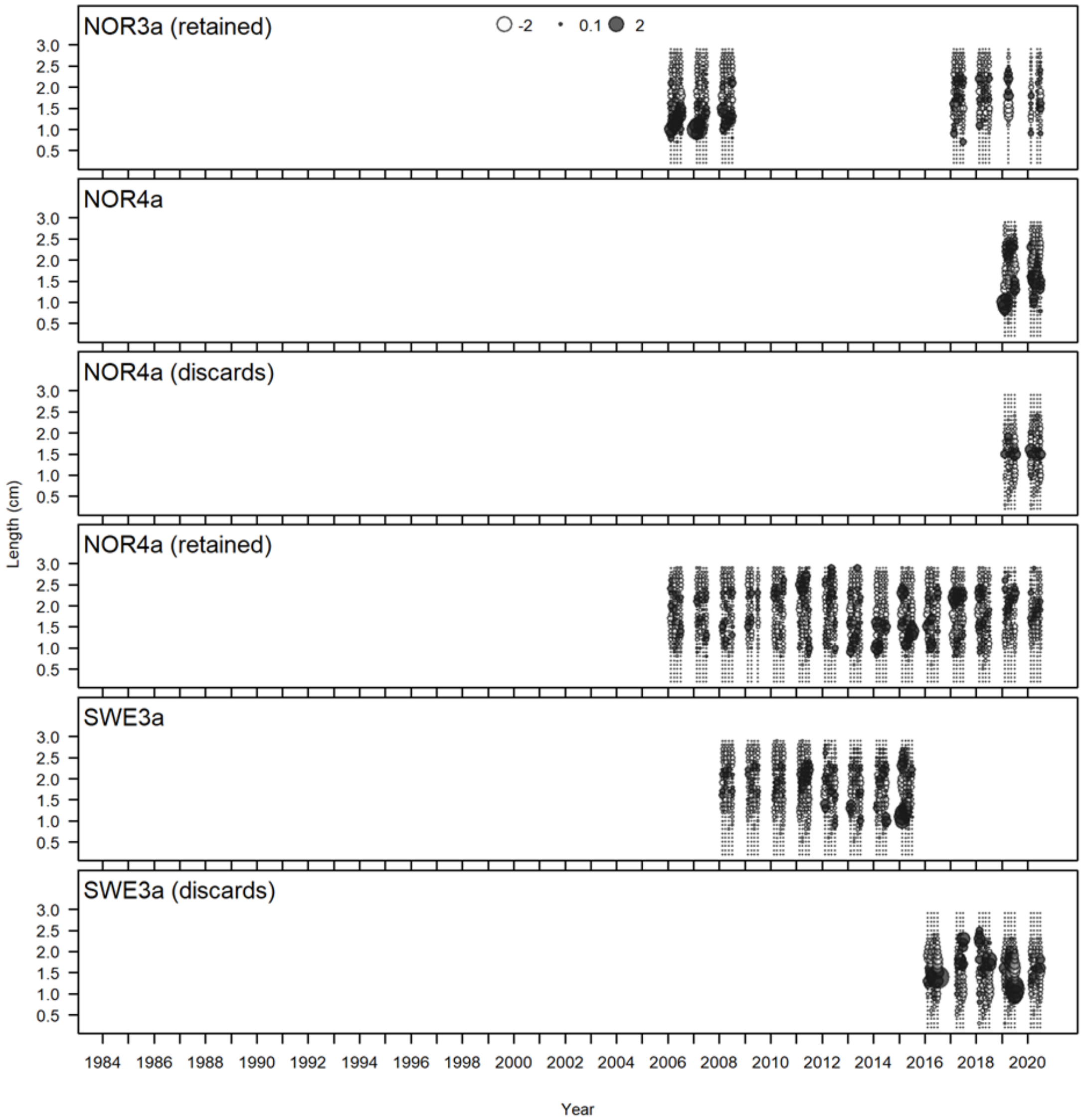


Figure 3.4.13 (continued). Northern shrimp in divisions 3.a and 4.a East. Residuals of fits to LFD data for the different fleets.

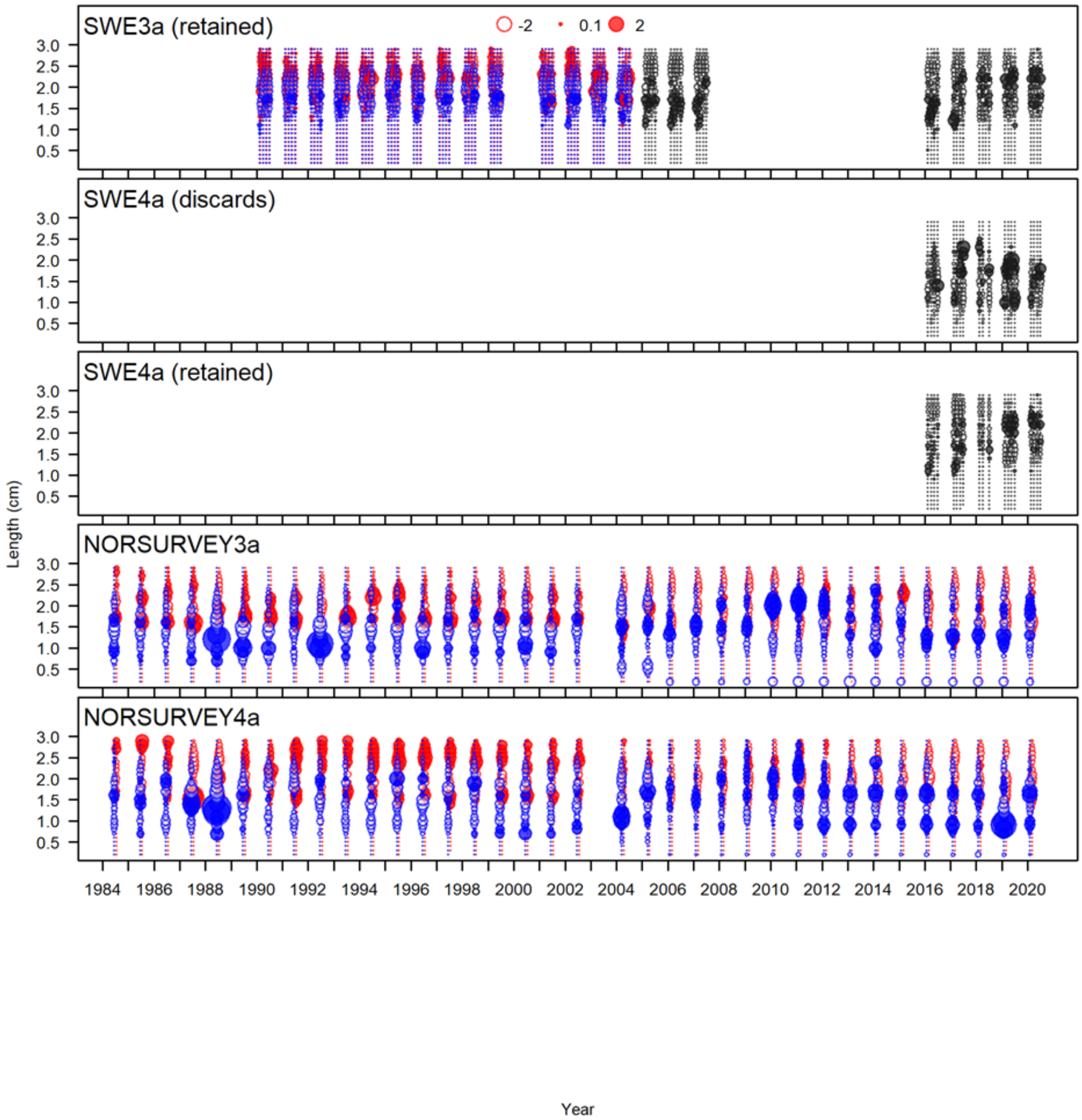


Figure 3.4.13 (continued). Northern shrimp in divisions 3.a and 4.a East. Residuals of fits to LFD data for the different fleets.

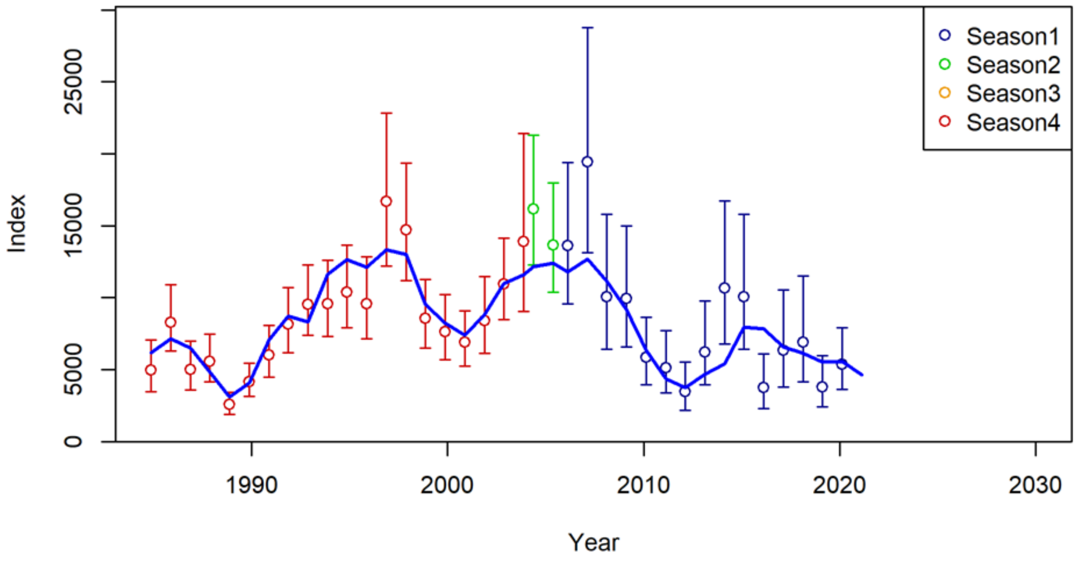


Figure 3.4.14. Northern shrimp in divisions 3.a and 4.a East. Model fits to the survey index in area 3a.

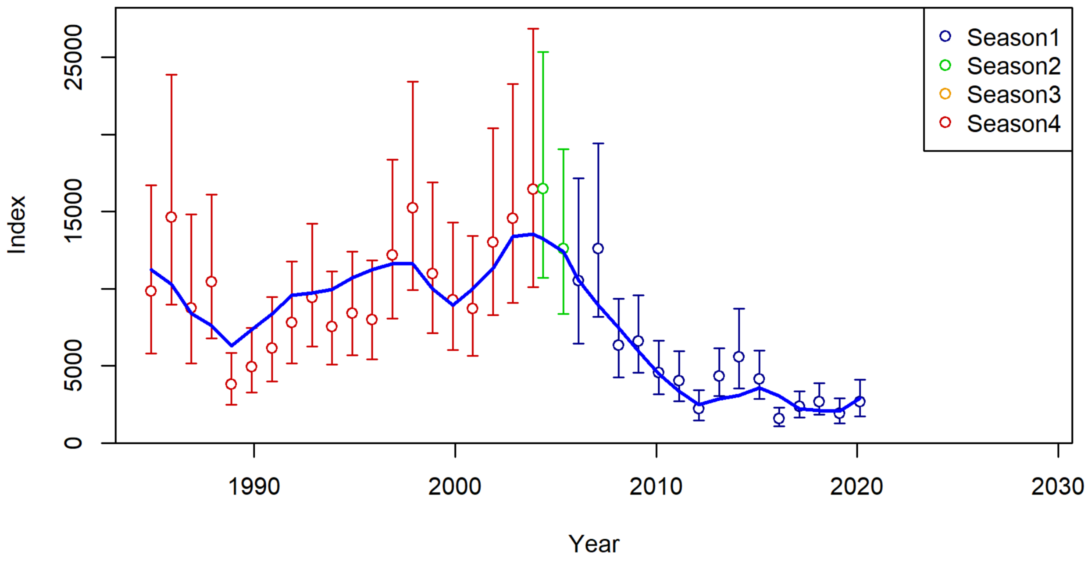


Figure 3.4.15. Northern shrimp in divisions 3.a and 4.a East. Model fits to the survey index in area 4a.

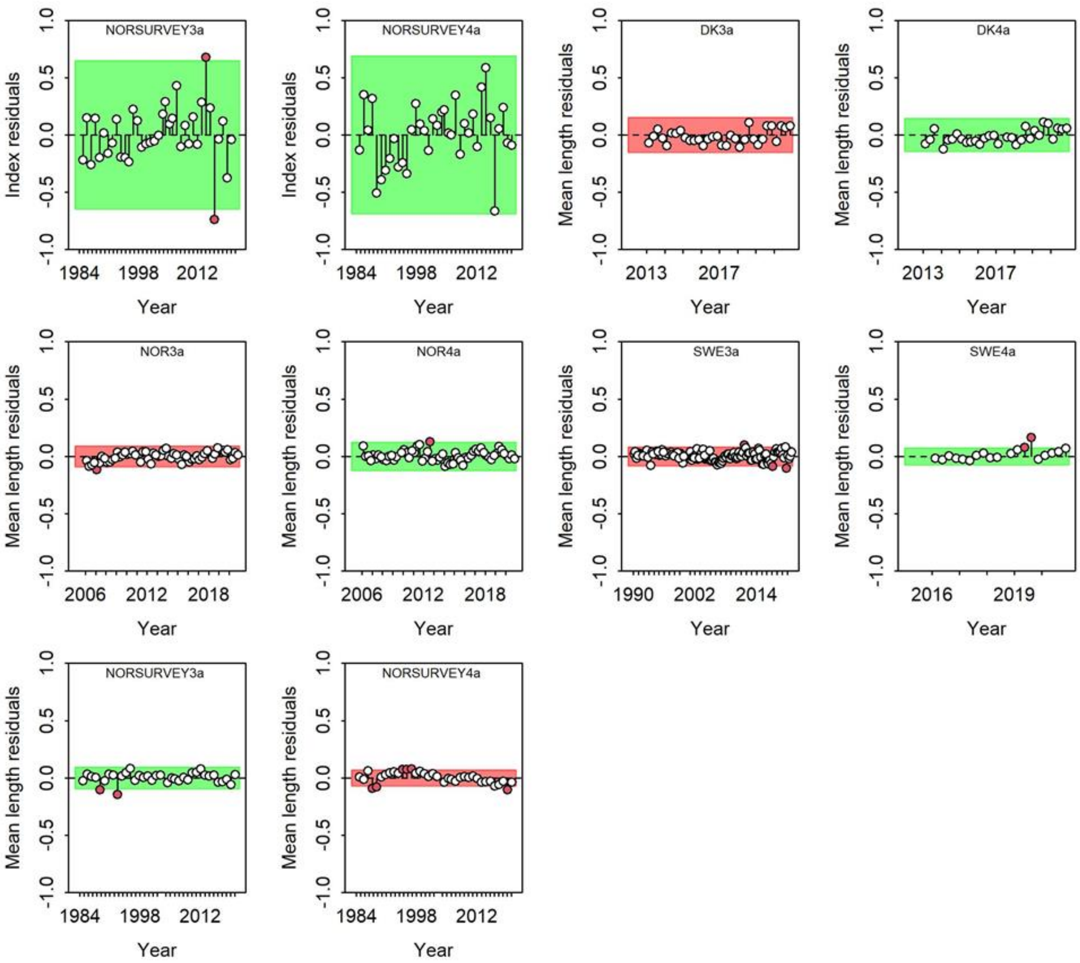


Figure 3.4.16. Northern shrimp in divisions 3.a and 4.a East. Residuals from runs test analyses for the fit to the survey indices and LFDs of commercial fleets and surveys.

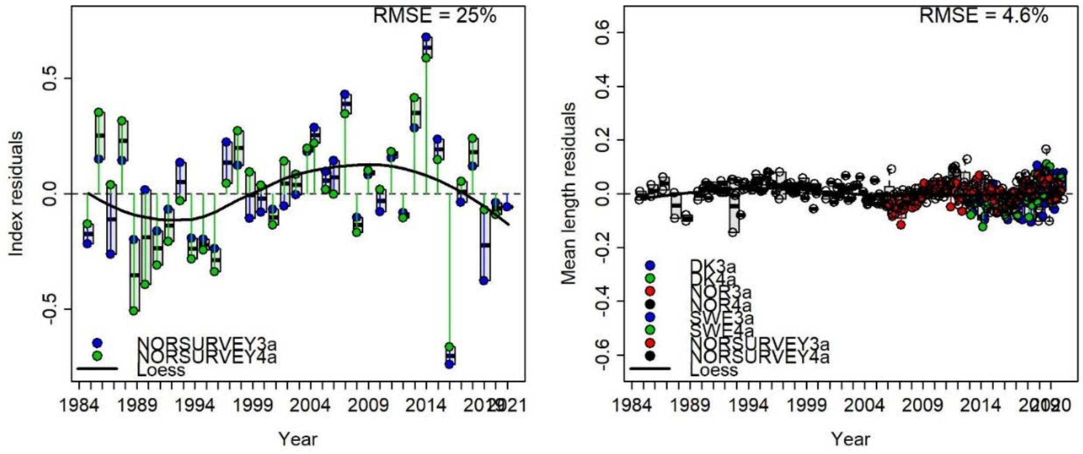


Figure 3.4.17. Northern shrimp in divisions 3.a and 4.a East. Residuals from the RMSE runs test analyses for the LFDs and the fit to the survey indices.

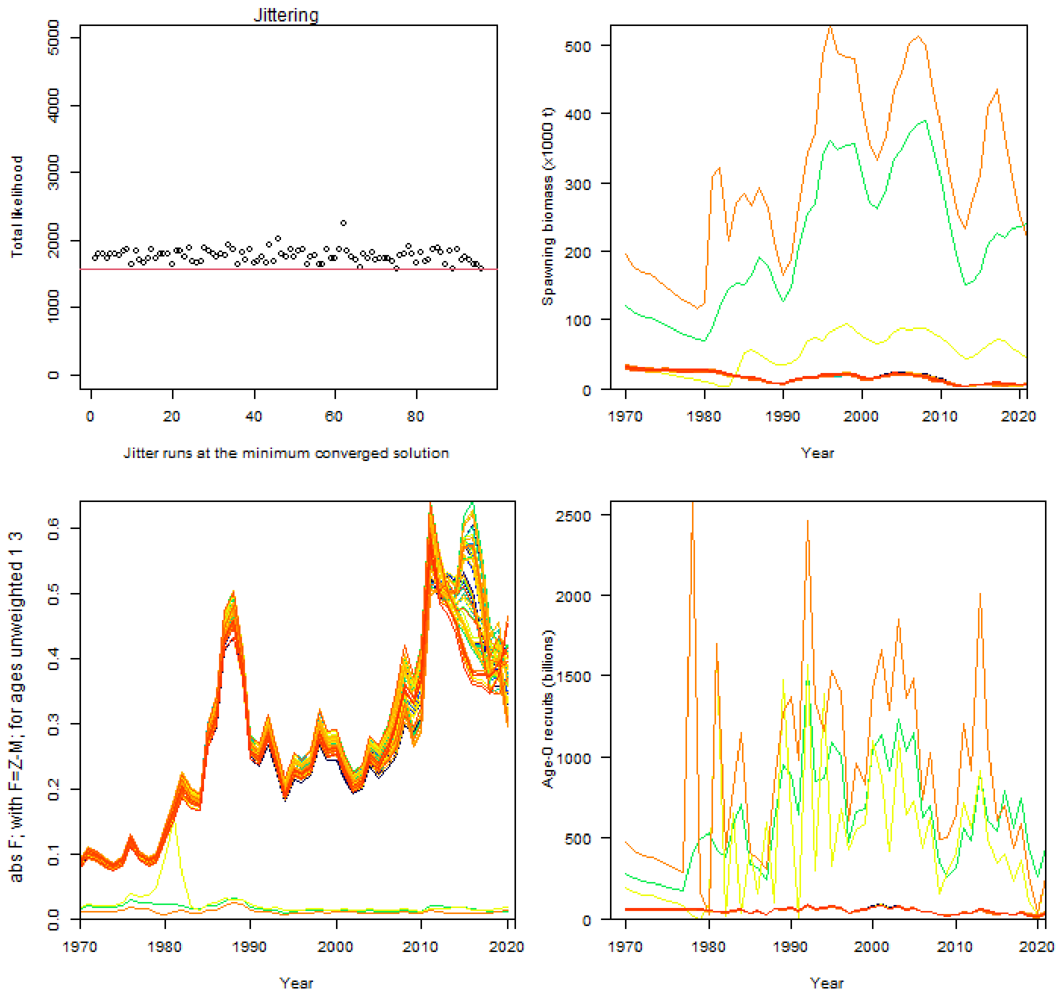


Figure 3.4.18. Northern shrimp in divisions 3.a and 4.a East. Results from jitter using 100 iterations and an average jitter of 10%.

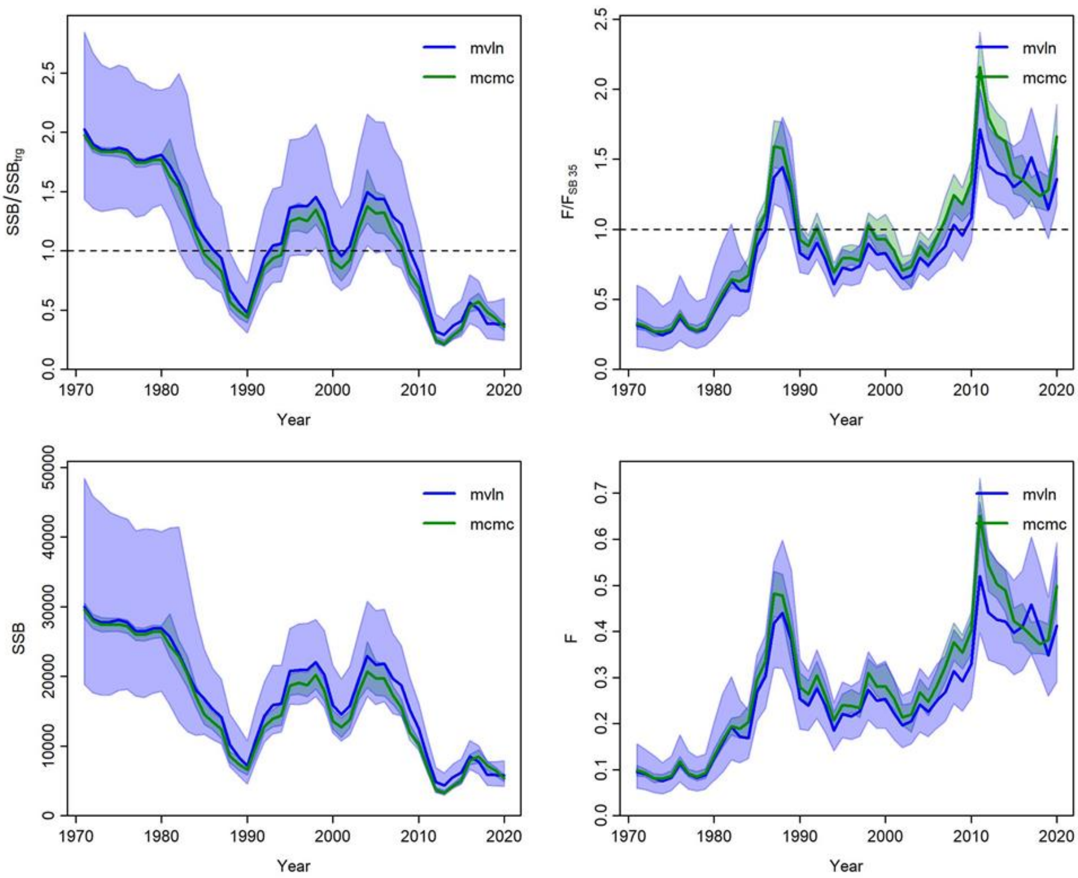


Figure 3.4.19. Northern shrimp in divisions 3.a and 4.a East. Results of the MCMC analysis in terms of SSB, R and F compared to the MLE model.

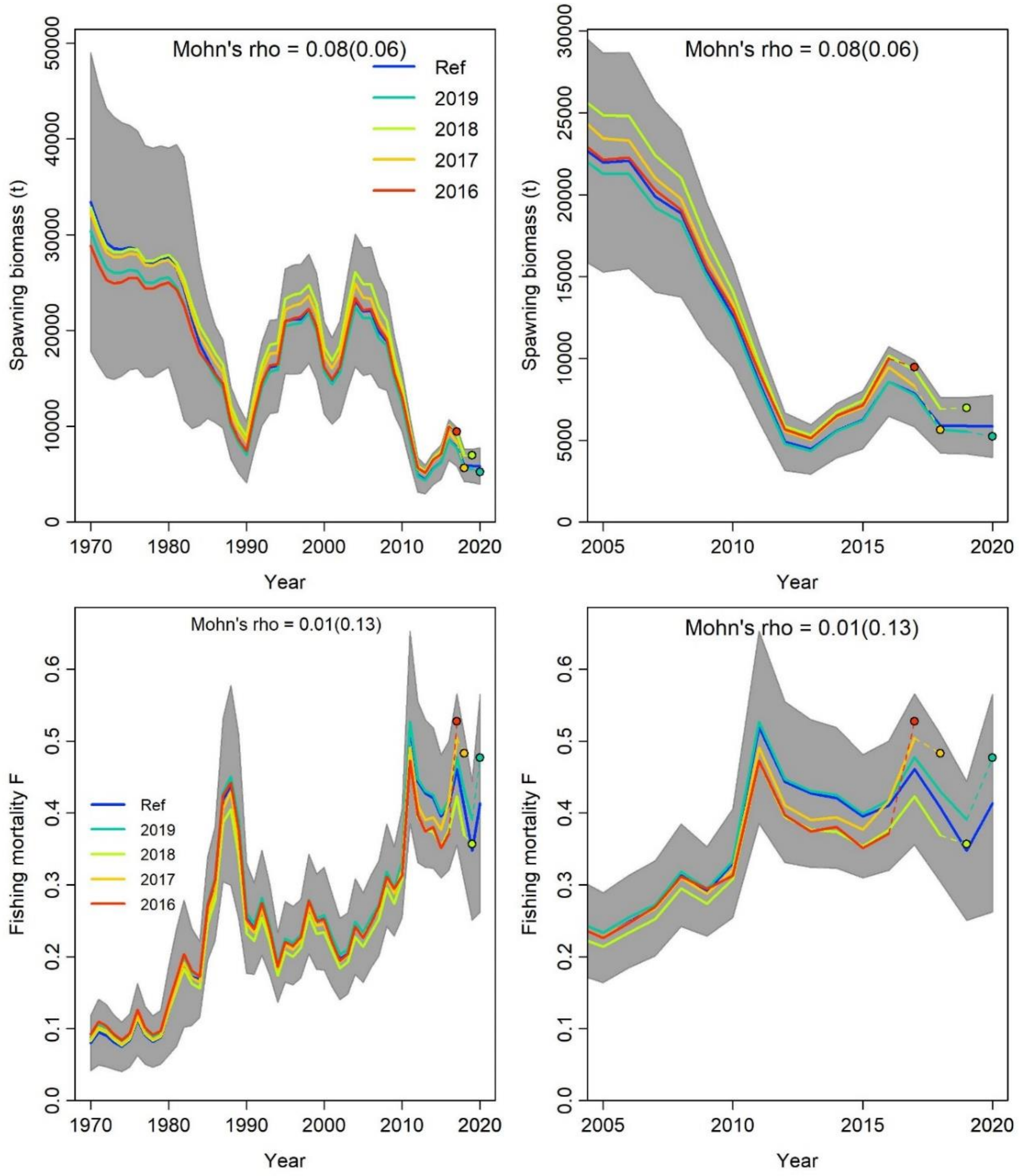


Figure 3.4.20. Northern shrimp in divisions 3.a and 4.a East. Retrospective analyses of the final model.

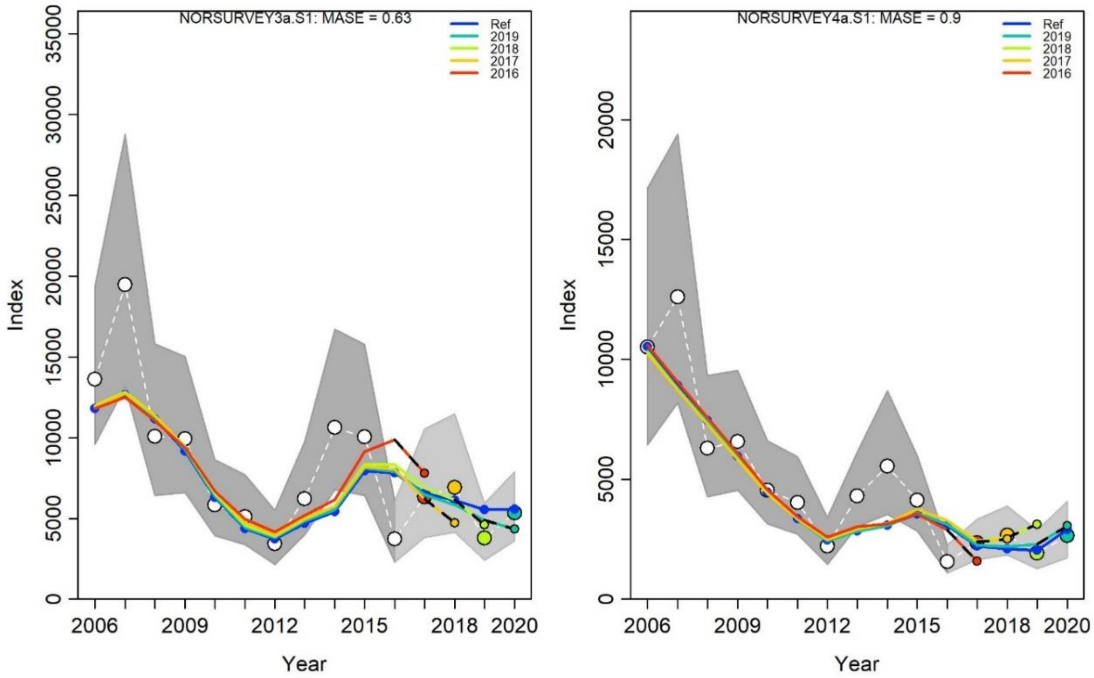


Figure 3.4.21. Northern shrimp in divisions 3.a and 4.a East. Hindcasting results for surveys showing observed (large white points connected with dashed line), fitted (solid lines) and one year-ahead forecast values (small terminal points). HCxval was performed using one reference model (Ref equal to last year data 2020) and 4 hindcast runs (solid-coloured lines) relative to the expected index. The observations used for cross-validation are highlighted as color-coded solid circles with associated 95 % confidence intervals (light-grey shading). The mean absolute scaled error (MASE) score associated with the survey index is denoted in each upper part of the panel

Table 3.4.9. Northern shrimp in divisions 3.a and 4.a East. Hindcasting results of the LFDs of the commercial and surveys and of the survey indices.

LFD	Season	MASE.adj
DK3a	1	0.79
DK4a	1	0.64
NOR3a	1	0.14
NOR4a	1	0.29
SWE3a	1	0.42
SWE4a	1	0.20
NORSURVEY3a	1	0.95
NORSURVEY4a	1	0.74
joint		0.52

LFD	Season	MASE.adj
DK3a	2	0.69
DK4a	2	0.20
NOR3a	2	0.15
NOR4a	2	0.38
SWE3a	2	0.19
SWE4a	2	0.11
joint		0.32

Index	Season	MASE.adj
NORSURVEY3a	1	0.63
NORSURVEY4a	1	0.90
joint		0.75

LFD	Season	MASE.adj
DK3a	3	1.02
DK4a	3	0.60
NOR3a	3	0.46
NOR4a	3	0.46
SWE3a	3	0.37
SWE4a	3	0.61
joint		0.59

LFD	Season	MASE.adj
DK3a	4	0.91
DK4a	4	0.48
NOR3a	4	0.58
NOR4a	4	0.37
SWE3a	4	0.73
SWE4a	4	0.83
joint		0.65

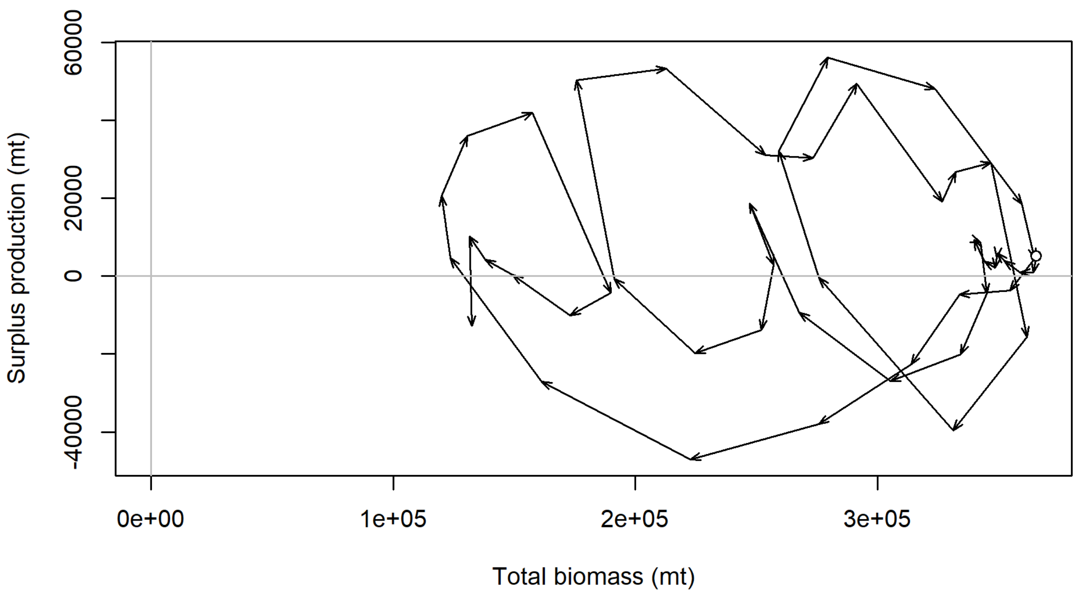


Figure 3.4.22. Northern shrimp in divisions 3.a and 4.a East. Surplus production against biomass plot. The round circle represents the first year of the time-series (1970).

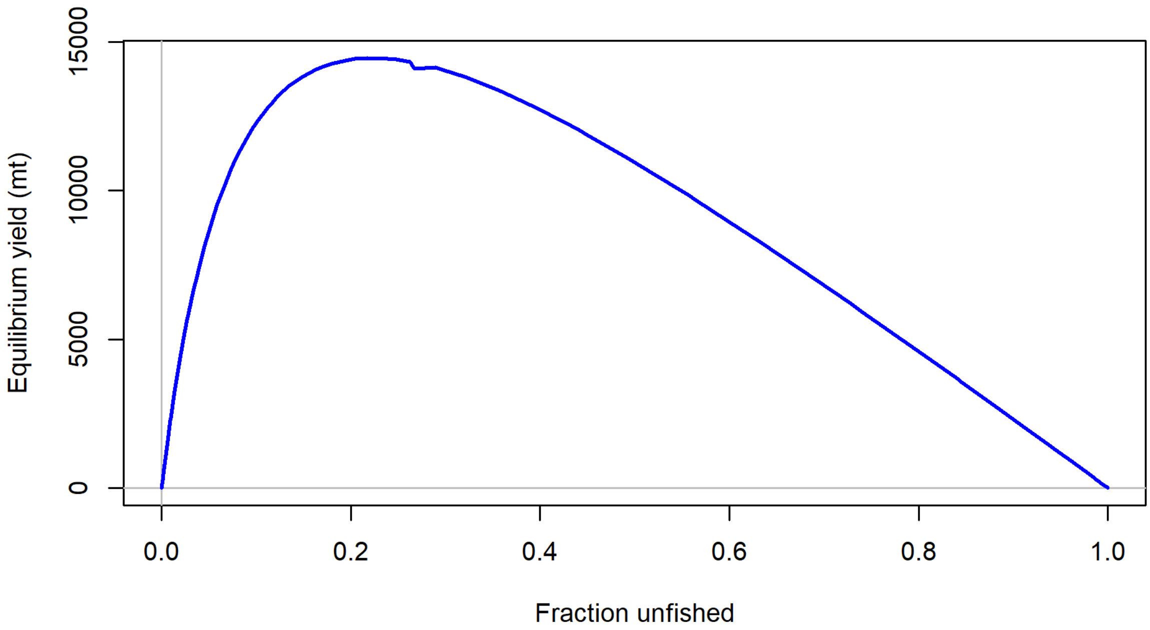


Figure 3.4.23. Northern shrimp in divisions 3.a and 4.a East. Production function.

Table 3.4.10. Northern shrimp in divisions 3.a and 4.a East. Summary table of the diagnostics of the final model and the same model without time varying discard ogive (base case described above). "Passed tests" score refers to the average test passes in % when multiple tests have been conducted.

Diagnostic	Indicator	Component	Reference	Discards ogive not time varying
Convergence		Model	0.0111692	2.21E-03
N. of parameters		Model	1323	1303
Hessian		Model	Yes	Yes
Jittering (10%)	% of runs under reference LL	Local minima		100%
	N of runs not different from reference run	Model		3%
Retrospective (4 years)	Mohn´s rho	SSB	0.08	-0.07
		F	0.01	0.12
Hindcasting	Forecast_Rho	SSB	0.06	0.35
		F	0.13	0.35
	MASE	Survey 3a	Pass	Pass
		Survey 4a	Pass	Fail
		Joint	Pass	Pass
		Lenght compositions joint	Pass	Pass
Run´s test		Survey	Pass	50%
		Lenght compositions	50%	38%
		RMSE survey	100%	100%
		RMSE Length compositions	100%	100%
		Passed tests	95.8%	65.7%

3.4.11 Ensemble model

The main input parameters of a stock assessment are often uncertain. This means that stock assessors are often faced with a range of model formulations and/or alternative management scenarios which should be scrutinized before decisions are made (Mannini *et al.*, 2021). In this context, when discussing which could be the best model used in assessing stocks, Hilborn and Walters (1992) recalled an adage that "*the truth often lies at the intersection of competing lies*". This uncertainty in 'what is the best model?' necessitates a comparison of a range of alternative models.

The biggest novelty used in this benchmark assessment is that, instead of comparing multiple model outputs and selecting a single final model, an ensemble modelling approach (Dietterich, 2000) was used. Ensemble methods provide a promising approach when decisions must be made despite the presence of multiple and potentially conflicting estimates of stock status (Anderson *et al.*, 2017). Ensemble models have been proven to be more accurate and less biased than the choice of an individual model, as they can effectively tease apart the conditions under which various model assumptions result in the most accurate predictions. In general, an ensemble approach will better encapsulate the variability and uncertainty of model predictions because instead of choosing a single set of fixed parameter values, you can explore a contrasting but plausible range of values. (Dietterich, 2000; Knutti *et al.*, 2009). This is crucial when the reliability of single fixed parameters is in question. The objective when using an ensemble model is therefore to quantify the total uncertainty across all plausible models, where the structural uncertainty is likely to be much greater than the within model uncertainty. For example, ensembles are often helpful because modellers need not decide on dome versus asymptotic fisheries selectivity (e.g. Sampson & Scott, 2012, FAO-GFCM, 2021), or whether to fix or estimate natural mortality (e.g. Johnson *et al.*, 2015).

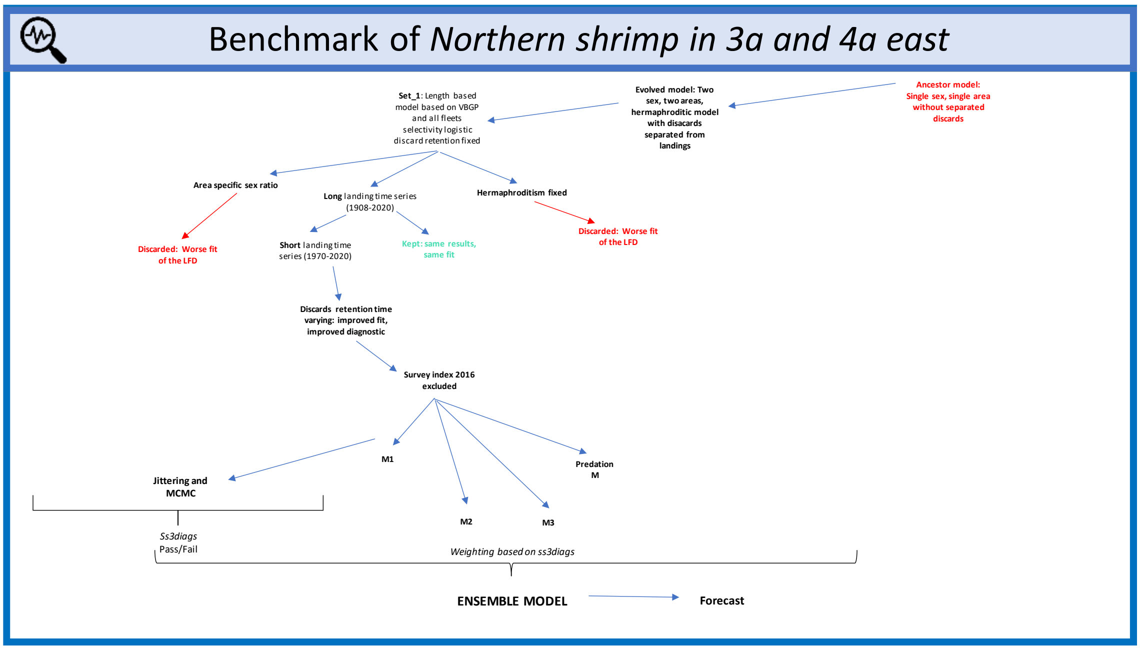
Natural mortality (M) is often considered as one of the most difficult to estimate, yet most influential parameters in stock assessment models (Mannini *et al.*, 2021). Here, the ensemble approach is used to incorporate three different, but equally plausible scenarios for M for the *Pandalus* stock in 3.a and 4.a East. In short, this involves running the final model (described in section 3.4.10) three separate times with three different M values (ensemble model grid listed in Table 3.4.11), whereby each model is considered a plausible 'state of nature' of the stock. The input files for each of these runs is available on the official ICES SharePoint (https://community.ices.dk/Expert-Groups/benchmarks/2022/WKPRAWN/2021_Meeting_Documents/06.%20Data/Pand_SKND/Ensemble.zip).

Table 3.4.11. Northern shrimp in division 3.a and 4.a East. Parameters and levels used in the ensemble model assessment grid. Details of the three M values are provided in section 3.4.2.4.

Parameter	Levels	Pregressive number of runs	Values
Natural mortality *(M)*	3	3	Lower M: CI 5% combine M; Medium M: Median combine M; Upper M: CI 80% combine M;

A schematic graphical representation of the assessment workflow is provided in Figure 3.4.24. Its inclusion is designed to provide a guideline via which the process of ensemble model grid construction can be followed as well as the steps taken prior to its implementation.

Figure 3.4.24. Northern shrimp in division 3.a and 4.a East. Schematic graphical representation of the assessment work-flow.



Model weighting

Before running the ensemble model, each plausible model needs to be assigned a weight. The need to weigh models based on available information is well recognised (Francis and Hilborn, 2011), but it is often difficult to do so within the context of fisheries stock assessment models as their complexity prevents strict adherence with statistical rigor. Despite is, assigning weights is a necessary step because assigning the same weight (reliability) to all models could introduce biases into the management advice if some models are, in fact, highly unlikely. To assign weights to the various models (and their associated M scenarios), it is preferable to establish a system of discrete weight categories. Here, we decided to use diagnostic scores (*W(Diagnostics)*) as weighting metrics (Maunder *et al.*, 2020) to judge the plausibility of each model based on its fit to the data. In this context, the *W(Diagnostics)* component is calculated based on a series of inter-connected diagnostic tests as discussed by Carvalho *et al.* (2021) and previously presented and explained in section 3.4.10 for the final model:

$$W(Diagnostics)\text{:}\ \frac{\boldsymbol{W}(\textbf{Diags 1}) + \boldsymbol{W}(\textbf{Diags 2}) + \boldsymbol{W}(\textbf{Diags 3}) \ldots + \boldsymbol{W}(\textbf{Diags N})}{\textbf{Num of } \boldsymbol{W}(\textbf{Diags})}$$

where each W component is assigned a value of 1 when the run passes the diagnostic test and a 0 when it fails. A summary of all main diagnostics for the three model runs is provided is Table 3.4.12. Based on these results, different weights were assigned. For clarity, the assigned weights are used as scaling factors for the number of simulations in the ensemble model, where a model with a weight of 100% is given 5000 simulations and a model with weight of 50% is given 2500 simulations.

Table 3.4.12. Northern shrimp division 3.a and 4.a *E*ast. Summary table of the diagnostics used in the weighting procedure. Green refers to a "Passed" score.

	Convergence and stability		Goodness of the fit												Consistency			
	Positive	Jittering	Run test												Retrospective analysis			
Run name	Hessian		CPUE1	CPUE2	Len1	Len2	Len3	Len4	Len5	Len6	Len7	Len8	Index	Length	Retro_SSB	Forecast_SSB	Retro_F	Forecast_F
Run1	Passed	Passed	Passed	Passed	Failed	Passed	Failed	Passed	Failed	Passed	Passed	Failed	Passed	Passed	Passed	Passed	Passed	Passed
Run2	Passed		Passed	Failed	Failed	Passed	Failed	Passed	Failed	Passed	Passed	Failed	Passed	Passed	Passed	Passed	Passed	Failed
Run3	Passed		Passed	Passed	Passed	Passed	Failed	Passed	Failed	Failed	Passed	Failed	Passed	Passed	Passed	Passed	Passed	Passed

Prediction skills													
Hindcasting (MASE)													
Survey3a	Surve4a	Joint	Len3a	Len4a	lenA1S1	lenA1S2	lenA1S3	lenA1S4	lenA2S1	lenA2S2	lenA2S3	lenA2S4	Weight
Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	0.83
Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Failed	Passed	Failed	Passed	0.72
Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed	0.86

Running the ensemble model

Once all plausible models have been run and have been assigned weights, a delta-Multivariate log-Normal estimator (delta-MVLN; Walter and Winker, 2019; Winker *et al.*, 2019) was used to run the ensemble model. During this, the delta-MVLN generates and stitches together the joint posterior distributions of the target derived quantities (e.g. SSB/SSB_{target} and F/F_{target}). These quantities are derived by using the delta-method to calculate asymptotic variance estimates from the inverted Hessian matrix of the Stock Synthesis model (i.e. the quantities are calculated from each of the three model runs). The delta-MVLN is used to run the ensemble because it can infer within-model uncertainty from maximum likelihood estimates (MLEs), standard errors (SEs) and the correlation of the untransformed quantities. Moreover, the delta-MVLN has been demonstrated to mimic the Markov Chain Monte Carlo (MCMC) approach closely (Winker *et al.*, 2019) and is therefore suitable for the task.

Ensemble results

To recap, to better capture uncertainties around the value of M for the *Pandalus* stock in 3.a and 4.a East, three plausible models were selected, run, assigned weights using diagnostics (interconnected diagnostic tests; Carvalho *et al.*, 2021; Maunder *et al.*, 2020; Kell *et al.*, 2021), and stitched together in an ensemble using the delta-MVLN (Walter and Winker 2019; Winker *et al.*, 2019). The run specifications and final weighting factors used in the ensemble are reported in Table 3.4.13. The final outputs from the ensemble model are based on the weighted-median value of the three runs.

Table 3.4.13. Northern shrimp in division 3.a and 4.a East. Run specifications ad final weighting factors used in the ensemble model. Run 1 is equivalent to the final model for *Pandalus* in 3.a and 4.a East described in section 3.4.10

Name	Natural Mortality	Weighting
*run1**	M medium	0.86
run2	M lower	0.72
run3	M upper	0.86

Figure 3.4.25 and 3.4.26 present the main outputs from the ensemble model, whereby Figure 3.4.25 provides a comparison between the ensemble and the three separate model runs. The main trends from the ensemble are summarized as:

- <u>State of the adult biomass (*SBB*):</u> Total spawning biomass of *Pandalus* has declined from the beginning of 2000s to the latest years. The last estimate of *SSB* in 2020 is 5796 tonnes (95% CI: 3966–9549).
- <u>State of exploitation (*F*):</u> Fishing mortality is defined as the average *F* of age classes 1 to 3. Since the beginning of 2000s F has generally increased with a peak in 2012, remaining at high levels thereafter. The last estimate of *F* in 2020 is 0.41 (95% CI: 0.24–0.63)
- <u>State of the juveniles (*Rec*):</u> With the exception of the 2013-year class, recruitment has also shown a general declining trend within the last decade; last year's estimate of recruits (in 1000s) is 10 864 880 (95% CI: 4 057 177–31 891 496).

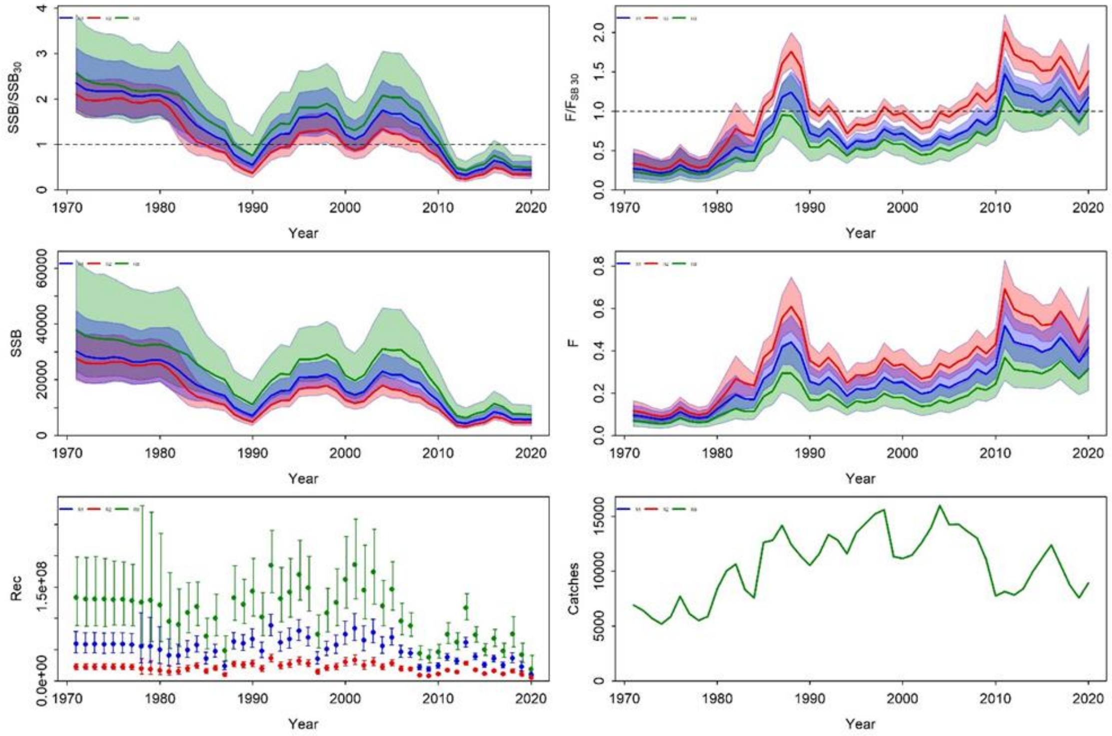


Figure 3.4.25. Northern shrimp division 3.a and 4.a East. Comparison of stock assessment results across the 3 single *model* runs (3 panels on the left) and the final ensemble model (3 panels on the right). Weighted-median value of SSB, F and Rec with 95% confidence intervals from delta-MVLN.

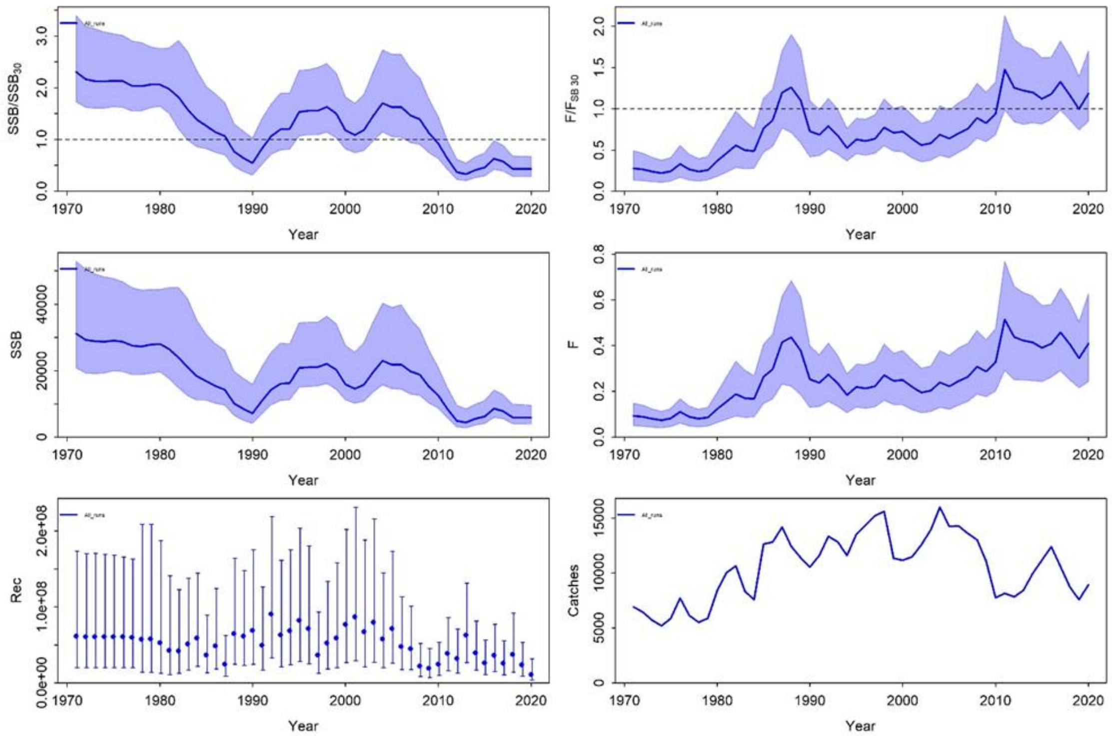


Figure 3.4.26. Northern shrimp division 3.a and 4.a East. Trajectory of the stock compared to the reference points B30%. Reference points are detailed in section 3.6.

A Kobe plot for the ensemble model in presented in Figure 3.4.27. The Kobe plot considers the time series of pressure (F/Ftarget) on the y-axis and the state of the stock's biomass (SSB/SSB_{target}) on the x-axis. The orange area indicates healthy stock sizes that are about to be depleted by overfishing. The red area indicates ongoing overfishing and that the stock is too small to produce maximum sustainable yields. The yellow area indicates that the biomass is too small/still recovering and that a reduction in fishing pressure is needed. The green area is the target area for management, indicating sustainable fishing pressure and a healthy stock size capable of producing high yields close to the chosen reference points (MSY or proxies).

The stock trajectory began in 1970 in the green quadrant, when the biomass was higher compared to the reference points (see section 3.6 for further details). In the period 1970–1990, the F level increased which resulted in a progressive erosion of the stock size, moving the stock trajectory towards the yellow quadrant. Following this, F remained below the reference points for the next 20 years (1990-2010), allowing a partial recovery of the biomass. Since 2010 a sharp increase in F and a subsequent stabilization brought the stock into a state of over-exploitation. For this reason, over last 10 years the stock has been in the red quadrant of the plot. In 2020 there was an approx. 83% probability that the stock is in the red quadrant of the Kobe plot (i.e. SSB > SSB_{30} and F < F_{30}) with lower probability (approx. 17%) of being in the yellow (i.e. SSB < SSB_{30} and F < F_{30}) and almost 0% probability of being in the green (SSB < SSB_{30} and F > F_{30}).

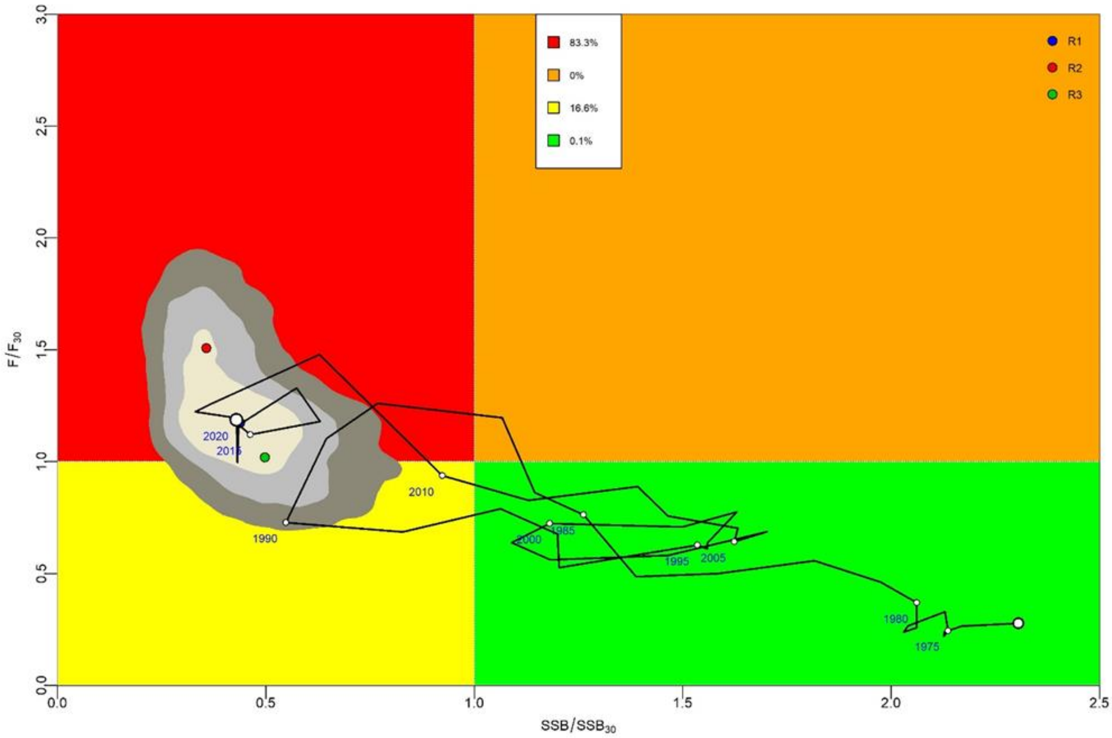


Figure 3.4.27. Northern shrimp division 3.a and 4.a East. Kobe plot showing the trajectory of relative stock size (SSB/SSB30) over relative exploitation (F/F30) based on the final ensemble model (white dot: weighted-median value of 3 runs). Grey shading indicates CI of 50%, 80% and 95% from delta-MVNL of the final assessment year (2020). The legend indicates the estimated probability of the stock status being in each of the Kobe quadrant.

The ensemble model and its results were presented to the WKPRAWN meeting and following some discussions it was proposed as the final model for providing scientific advice for *Pandalus* in 3.a and 4.a East.

3.5 Short-term projections

Settings for short-term projections, used to provide catch advice, were discussed at the meeting. In SS3 it is possible to forward project the population under a range of catch and F scenarios. Achieving biomass targets is more problematic but approximate values for F and catch can be interpolated for target biomass if required. The agreed settings are indicated in the Stock Annex.

3.6 MSE and Reference Points (MSY)

Management strategy evaluations (MSE) were used to determine the target and trigger reference points to be used to provide advice for *Pandalus* in 3.a and 4.a East. To this aim, we used the simulation-testing framework available in the Fisheries Library for R (FLR; Kell *et al.*, 2007; https://flr-project.org/). The simulation framework was implemented in the FLR library \`mse\` (https://github.com/flr/mse) with \`FLasher\` (https://github.com/flr/FLasher) being used to carry out the forward projections. Reference points at equilibrium were calculated with \`FLBRP\` (https://github.com/flr/FLSRTMB). To facilitate customized reference point estimation and visualization of F_{MSY} proxy (hereafter defined as F_{brp}, which in this case was expressed as the F that brings the stock at a given fraction of B_0, i.e. FB%), B_{lim}, FP.05, B_{trg}, F_{trg}, we developed the FLR package \`FLRef\` (https://github.com/henning-winker/FLRef). \`FLRef\` makes use of the new fast forward projection 'ffwd()\` in \`FLasher\` together with the bisection function \`bisect()\` in 'mse' to efficiently derive precise values of FP0.5 based stochastic simulations. R code used in this analysis will be made available in the GitHub repository of \`FLRef\` and in the ICES SharePoint of WKPRAWN.

MSE simulations were run for the three models included in the ensemble (hereafter defined as R1, R2 and R3). The operating models were implemented as single sex and single fleet models with an annual time step. Future projections were run for 60 years (i.e. 2021-2080) with 250 iterations, and were based on the 3-year average of the most recent data (i.e. 2017-2020) for weight-at-age, maturity-at-age, natural mortality-at-age and the F pattern determining the selectivity-at-age. This choice was made to account for non-stationary processes in these quantities. The performance evaluations were based on the last 10 years of the 60-year projection horizon (i.e. 2071-2080). For the simulation testing, stock and recruitment, steepness, sigma R and autocorrelation were all set at the same values previously derived for each model of the ensemble. The recruitment deviation is assumed to be associated with a first-order autocorrelation (AR1) process and a function of recruitment standard deviation σ and the AR1 coefficient ρ (Johnson *et al.*, 2016). Simulations included implementation error as estimated using the last 6 years (average +8%, sd = 0.12; ICES, 2021b).

Harvest control rules (HCRs) were kept generic and in the same form of the conventional ICES Advice Rule (ICES, 2021a), where the advice decreases from F_{trg} to zero and from $B_{trigger}$ to zero SSB. Variations of the tested HCRs are therefore determined by the parameters F_{trg} and $B_{trigger}$. The HCRs were implemented using a simulated feedback control loop between the implementation system and the operating model, where the implementation system translates the assessment outcome via the HRC into the Total Allowable Catch (TAC) advice (Figure 3.4.28). The key difference to a simple stochastic risk simulation, such as EQsim, is the simulated feedback control loop between the implementation system and the operating model accounts for the lag between the last year of data used in the assessment and the implementation year of TAC advice. In ICES, the implementation system of HCR is based on the assumption that advice is given for year y+1 based on an assessment completed in year y, which is typically fitted to data up until year y-1 (ICES, 2020). Therefore, implementation of the TAC derived through HCR requires projection of the stock dynamics by way of a short-term forecast (Mildenberger *et al.*, 2021). In contrast to a full MSE simulation design (Punt *et al.*, 2014), a MSE 'short-cut' approach (e.g. ICES, 2020) omits the step of annually updating the estimation model (assessment) in the feedback control. Instead, it passes the 'true' age-structured dynamics from the operating model (or with assumed error) to the HCR implementation. The merits of a short-cut MSE approach include the incorporation of the lag effect between data, assessment, and management implementation. The limitations of the MSE short-cut approach are that it cannot fully account for uncertainties resulting from imperfect sampling of the full age-structure (e.g. poorly sampled recruits), observation error and model estimation error.

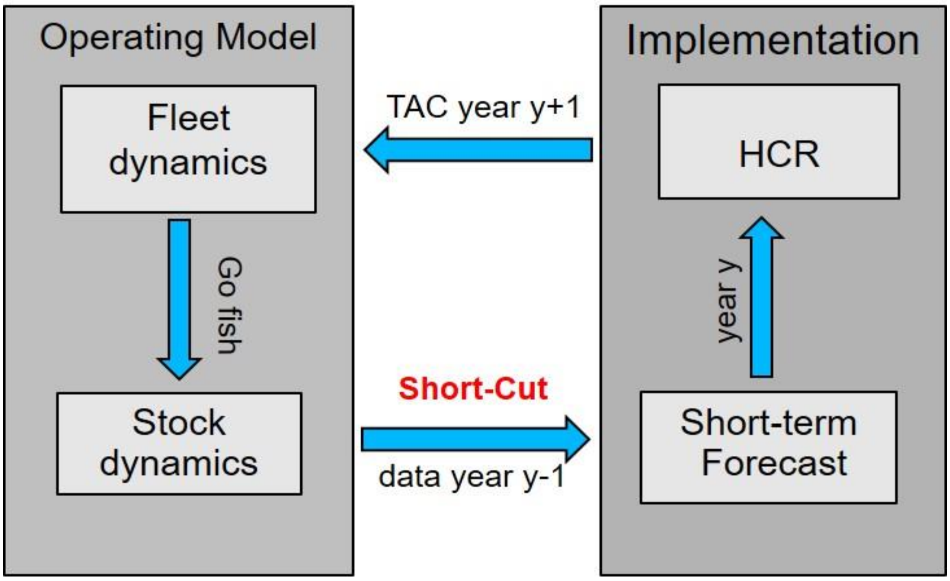


Figure 3.4.28. Northern shrimp in divisions 3.a and 4.a East. Schematic illustrating the key processes of the short-cut approach to MSE, showing the Operating Model that simulates the fishery and stock dynamics on the left and Implementation System including the short-term forecast on the right. The short-cut denotes the omission of the estimation (stock assessment) model which updates with new observations (with estimation error) in a conventional MSE implementation with a full feedback control loop.

Performance Evaluation Criteria

The consistency tests were designed to identify the generic rules for specifying F_{brp}, B_{trg} and $B_{trigger}$ according to the stock-specific productivity that provides the optimal trade-off among the following three main objectives: (1) to not exceed a 5% probability of SSB falling below B_{lim} in any single year (2) to achieve high long-term yields that correspond to at least 95% of the median long-term yield attained by fishing at the deterministic F_{MSY} (MSY), (3) to attain at least 50% probability that SSB is above 80% of B_{trg}. Consistent with the objectives of the ICES advice framework (ICES, 2020), these three objectives are interpreted hierarchically whereby (1) is the overriding criteria of maintaining stock size above B_{lim} with at least 95% probability to be compliant with the ICES Precautionary Approach (PA). Conditional on objective (1), objective (2) is based on the ICES definition for using plausible values around F_{MSY} in the advice rule, which are derived so that they lead to no more than a 5% reduction of MSY obtained by fishing at F_{MSY} in the long term.

In the previous assessments (e.g. ICES, 2016), the lowest observed SSB (i.e. B_{loss}) was used to derive B_{lim}. In ensemble model, B_{loss} will be inherently different for the different model configurations and therefore fractions of B_{MSY} or B_0 are used (ICES, 2022). Here we have chosen to define B_{lim} as a fraction of B_0, which when compared to B_{MSY} has the advantage of being independent to selectivity. When expressed as a fraction of B_0, B_{lim} typically ranges from 0.1 to 0.2 of B_0 (ICES, 2022). For *Pandalus*, B_{lim} was set at 15% of B_0, which is approximately the average B_{loss} for the three models (B_{loss} range: 11-17% of B_0). As shown by WKREF1 (ICES, 2022), setting B_{lim} well under 10% of B_0 renders $F_{P.05}$ ineffective for most ICES stocks with or without the use of $B_{trigger}$ (ICES, 2022). This is also particularly important in the presence of the Allee effect (i.e. depensation) in exploited fish, which was identified to occur when the stock is below 15-25% of B_0 (Perälä and Kuparinen, 2017; Perälä *et al.*, 2021).

MSE Results

Twelve scenarios (i.e. 3 x F_{B0} x 4 x $B_{trigger}$) and the deterministic F_{MSY} were tested for the three models of the ensemble. The stock-recruit relationship for the three different models of the ensemble is shown in Figure 3.4.29 while the reference points for R1 is shown as an example in Figure 3.4.30. As an example of the realised simulations, trends in SSB, F, landings, and R for the different combinations of F_{target} and $B_{trigger}$ as compared to the deterministic F_{MSY} are shown for R1 in Figure 3.4.31.

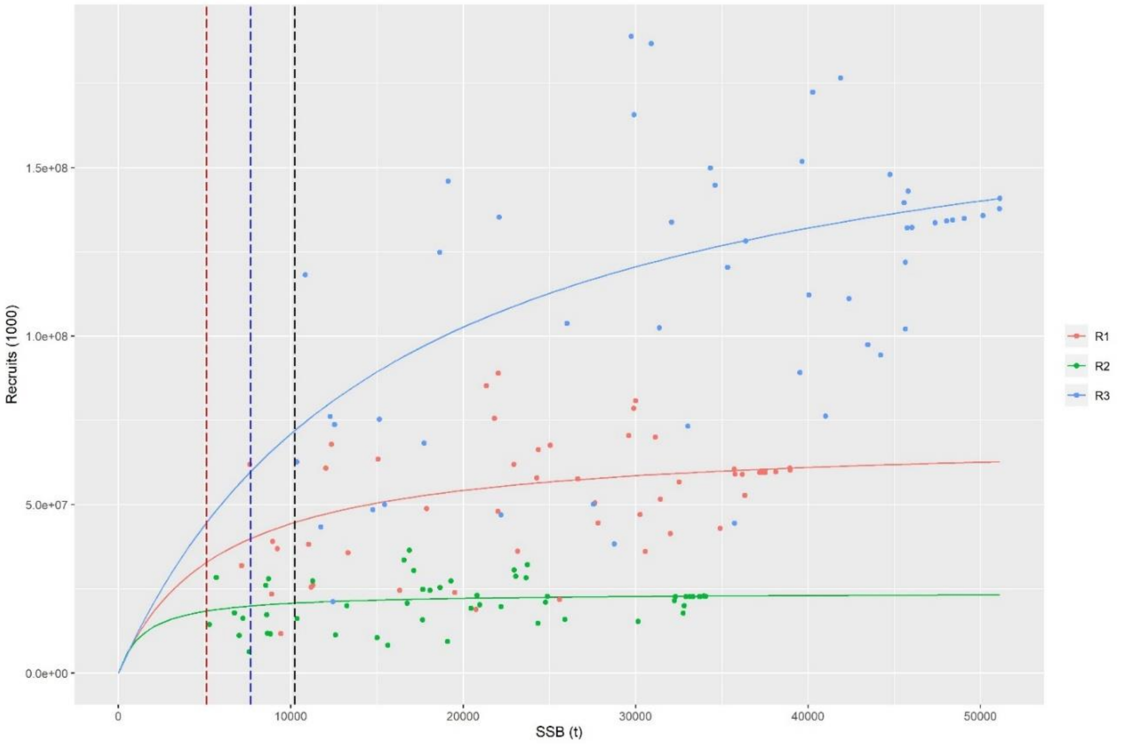


Figure 3.4.29. Northern shrimp in divisions 3.a and 4.a East. Stock-Recruitment relationship for the three models of the ensemble. Red, blue and black line are median $B_{10\%}$, $B_{15\%}$ and $B_{20\%}$ for the three models.

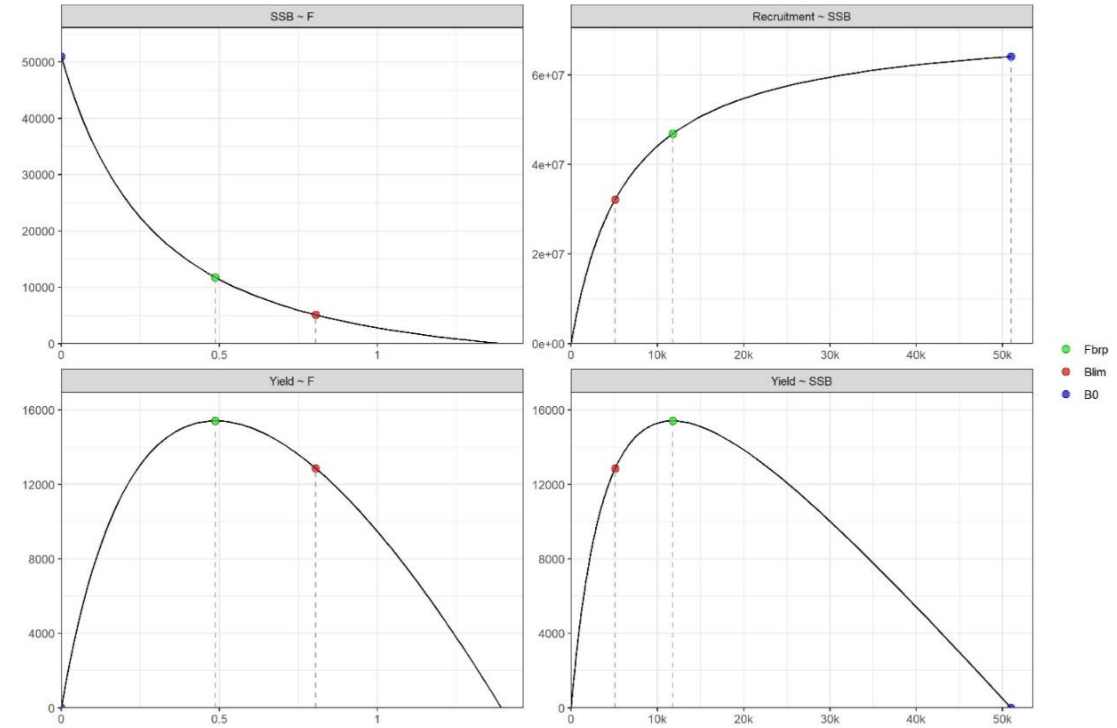


Figure 3.4.30. Northern shrimp in divisions 3.a and 4.a East. Estimated reference points for R1. B_{lim} is set as 15% of B_0, F_{brp} corresponds to $F_{B30\%}$.

The results of the MSE showed that $F_{B30\%}$ with $B_{trigger}$ set at 0.8 of $B_{30\%}$ achieve the highest long-term yields and have a median probability of SSB falling below B_{lim} which is less than 5% (Figure 3.4.32). Fishing at F_{MSY} is not precautionary as it implies a median probability of SSB falling below B_{lim} that is larger than 10% with extreme values up to 60%. The difference in long term yield between $F_{B30\%}$ with $B_{trigger}$ set at 0.8 and fishing at the determinist F_{MSY} is less than 3% with a long term SSB that is on average 30% larger than B_{MSY}. Moreover, by comparing $F_{B30\%}$ with $B_{trigger}$ set at 0.8 of $B_{30\%}$ and $F_{B30\%}$ with $B_{trigger}$ set at 0.9 of $B_{30\%}$ (Table 3.4.14) we show that they are almost equivalent when rounded to 2 decimal places. Thus, $F_{B30\%}$ with $B_{trigger}$ set at 0.8 of $B_{30\%}$ was accepted by the meeting.

For completeness, we run a bisect analysis for model R1 to estimate $F_{P.05}$, which is equivalent to EQsim. The results trajectories of SSB, R and catches with 1000 simulations are presented in Figure 3.4.33. The associated probability of falling below B_{lim} when fishing at $F_{B30\%}$ was 2.8%.

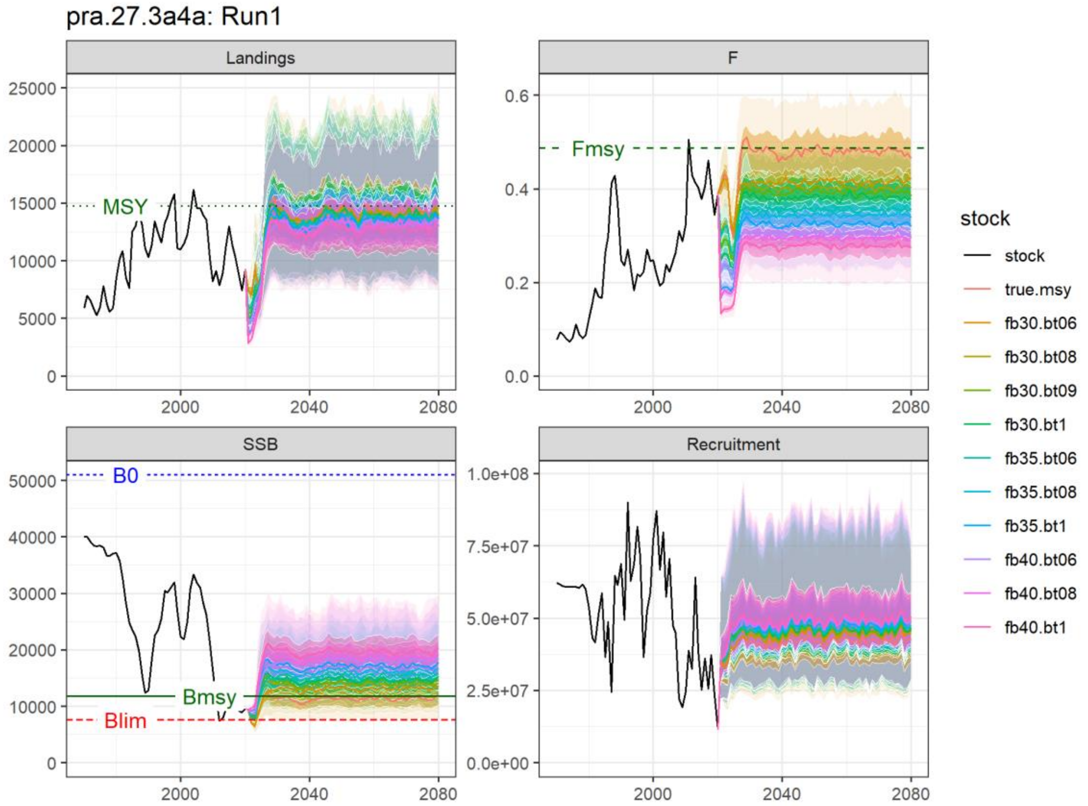


Figure 3.4.31. Northern shrimp in divisions 3.a and 4.a East. Long-term simulations for Run1. Trends in SSB, F, landings, and R for different combinations of F_{target} and $B_{trigger}$ and compared to the deterministic F_{MSY}.

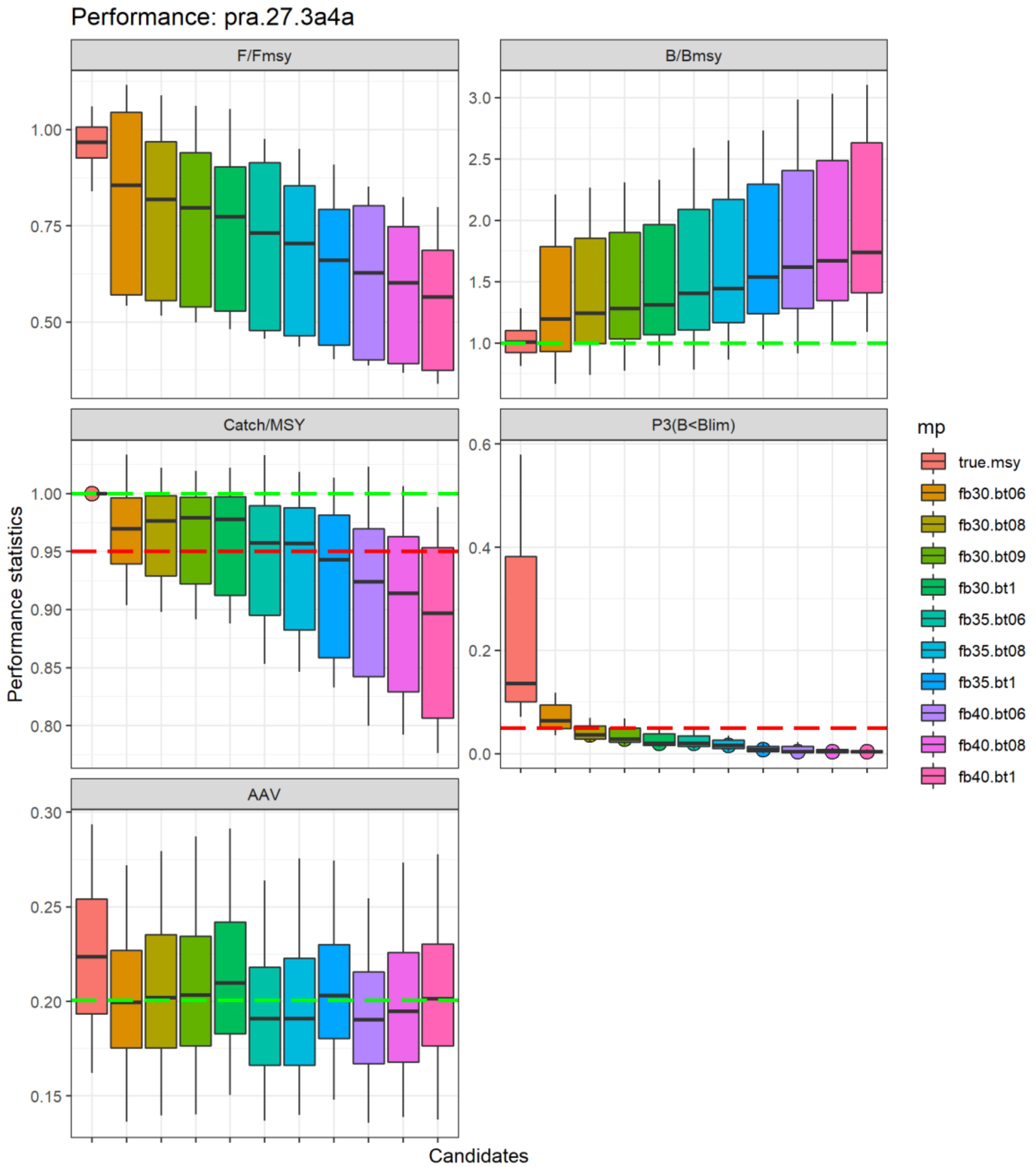


Figure 3.4.32. Northern shrimp in divisions 3.a and 4.a East. Results of the MSE used to evaluate reference point systems, showing the type 3 risk probabilities (P3) of SSB falling below B_{lim}, the median long-term yield relative the median long term obtained at fixed deterministic F_{MSY} (MSY), the median long-term F and SSB relative to the deterministic F_{MSY} and B_{MSY} and the median long term interannual variation in catches. Green and red dashed lines denoting the target and limit thresholds, respectively. Candidates based on $F_{B\%}$ and $B_{trigger}$ as fraction of $B_{\%}$.

Table 3.4.14. Northern shrimp in divisions 3.a and 4.a East. Selected results from the MSE.

Candidate	P3 (B < B_{lim})	Catch/MSY	B/B_{MSY}
Fb30.bt08	0.04	0.98	1.22
Fb30.bt09	0.03	0.98	1.25
Fb30.bt1	0.02	0.97	1.29

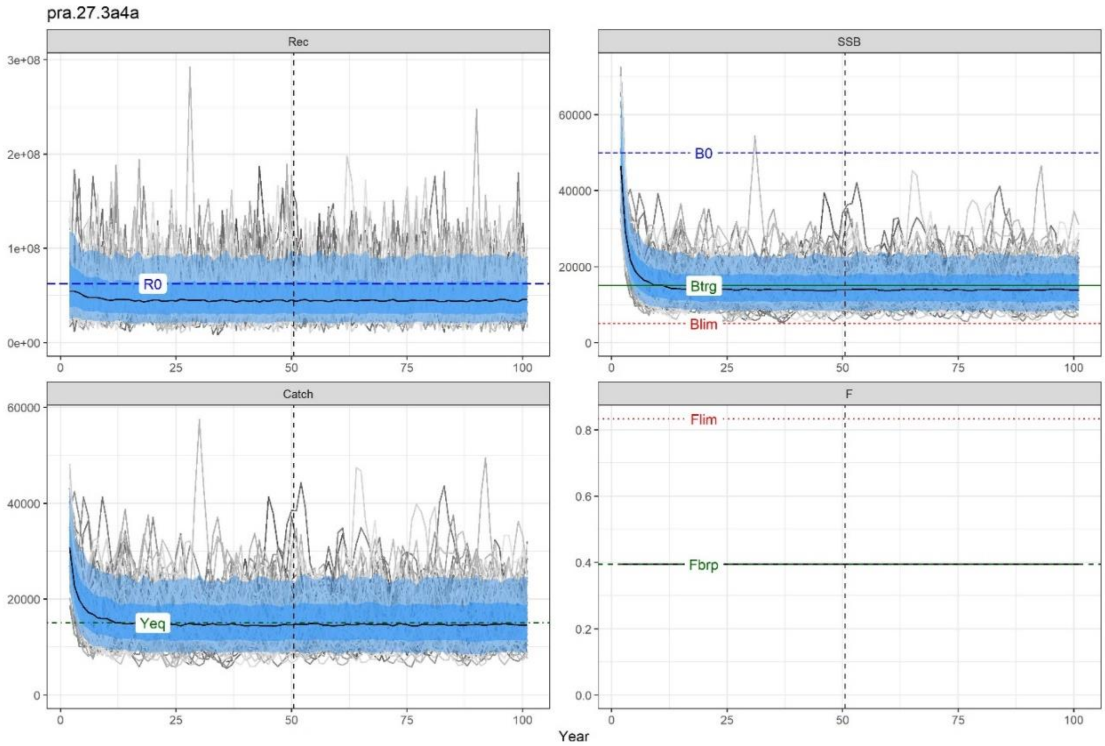


Figure 3.4.33. Northern shrimp in divisions 3.a and 4.a East. Results of the bisect analysis for model R1 to estimate $F_{P.05}$, which is equivalent to EQsim.

3.7 Future Research and data requirements

Live weight correction

Time and country varying fractions of *Pandalus* are landed with different presentations (i.e. boiled air cooled, boiled water cooled, iced for different durations and even frozen). The results to date indicate that this does lead to either weight loss or gain depending on the presentation type. This can lead to inaccuracies in catch statistics and potential biases in the way sampling data are worked up into total removals from the stock in numbers. WD11 summarises the results of different studies to date and this highlights that the issue is likely more complicated than first anticipated with missing historical data on presentation\onboard handling and further experiments needed. Differences in raising procedures (number vs. weight) were also identified as a potential issue. WKPRAWN recommends that conversion factors and raising procedures are investigated thoroughly before the next benchmark and this work should be coordinated with WGCATCH and involve the relevant *Pandalus* experts.

Time varying M predation model

WKPRAWN explored the utility of a predation index for this stock. The work done showed significant promise, however, further work is required before the approach can be implemented in assessment. First, an evaluation of uncertainties in the model is needed. Second, model simplification and data pruning should be evaluated to improve model fit and prediction. Finally, the composition of predators in the index needs to be settled. This work should be done intersessionally to feed into the next benchmark.

Weightings of ensemble

This was the first time and ensemble model has been used as the final assessment method by ICES. The approach taken to weight the individual runs within the ensemble described in section 3.4.11 should be considered further as the science and experience around ensemble modelling develops in ICES.

Sex change –currently age dependent

Within SS3 the sex change can only be modelled as an age dependent process. In the future it would be useful to have this length based but this is not currently possible.

Growth

The assessment diagnostic shows strong patterns in the residual fits to the length distributions. WKPRAWN suspected that these are most likely due to changes in growth or difference in growth across time and area not accounted for in the way growth is currently modelled. WKPRAWN recommended that this should be looked into further.

3.8 References

Anderson et al. 2020. An R package for fitting spatiotemporal species distribution GLMMs (generalized linear mixed effects models) with TMB (Template Model Builder) sdmTMB: (2020; PLOS Computational Biology) https://github.com/pbs-assess/sdmTMB

Berg, C. W., Nielsen, A., Kristensen, K. 2014. Evaluation of alternative age-based methods for estimating relative abundance from survey data in relation to assessment models. Fisheries Research 151: 91-99.

Breivik, O. N., Aanes, F., Søvik, G., Aglen, A., Mehl, S. and Johnsen, E. 2021. Predicting abundance indices in areas without coverage with a latent spatio-temporal Gaussian model. ICES Journal of Marine Science. doi:10.1093/icesjms/fsab073

Carvalho, F., Punt, A. E., Chang, Y. J., Maunder, M. N., & Piner, K. R. (2017). Can diagnostic tests help identify model misspecification in integrated stock assessments? Fisheries Research, 192, 28-40.

Carvalho, F., Winker, H., Courtney, D., Kell, L., Kapur, M., Cardinale, M., Schirripa, M., Kitakado, T., Ghebrehiwet, D.Y., Piner, K.R., Maunder, M.N., Methot, R., 2021. A Cookbook for Using Model Diagnostics in Integrated Stock Assessments. Fisheries Research, https://doi.org/10.1016/j.fishres.2021.105959.

Hatlebrekke, H.H., Gundersen, S., Nedreaas, K., Vølstad, J.H. and Kolding, J. 2021. The Coastal Reference Fleet 2007-2019. Fleet composition, fishing effort and contributions to science. Rapport fra Havforskningen 2021-52. ISSN:1893-4536. 61 pp.

Hjort, J. and Ruud, J. T. 1938. Rekefisket som naturhistorie og samfundssak. Fiskeridirektoratets skrifter. Serie Havundersøkelser Vol. V, No. 4. 158 pp. (in Norwegian).

ICES 2005. Report of the Pandalus assessment working group, 27 October – 5 November 2004. ICES C.M. 2005/ACFM:05, 74 p.

ICES. 2016. Report of the Benchmark Workshop on Pandalus borealis in Skagerrak and Norwegian Deep Sea (WKPAND), 20–22 January 2016, Bergen, Norway. ICES CM 2016/ACOM:39. 72 pp.

ICES. 2020. The third Workshop on Guidelines for Management Strategy Evaluations (WKGMSE3). ICES Scientific Reports. 2:116. 112 pp. http://doi.org/10.17895/ices.pub.7627.

ICES. 2021a. Benchmark Workshop on the development of MSY advice for category 3 stocks using Surplus Production Model in Continuous Time; SPiCT (WKMSYSPiCT). ICES Scientific Reports, 3: 1–317.

ICES. 2021b. Joint NAFO\ICES Pandalus Assessment Working Group (NIPAG). ICES Scientific Reports. 3:22. 25 pp. https://doi.org/10.17895/ices.pub.5990.

ICES. 2022. Workshop on ICES reference points (WKREF1). ICES Scientific Reports. 4:2. 70 pp. http://doi.org/10.17895/ices.pub.9749.

Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., & Ono, K. (2014). Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science, 72(1), 99-110.

Jørgensen, M. et al. 2014. Introducing time-varying natural mortality in the length-based assessment model for the Pandalus borealis stock in ICES Div. IIIa and IVa east.

Johnsen, E., Totland, A., Skålevik, Å., Holmin, A.J., Dingsør, G.E., Fuglebakk, E. and Handegard, N.O. 2019. StoX: An open source software for marine survey analyses. Methods in Ecology and Evolution 10: 1523-1528. DOI: 10.1111/2041-210X.13250

Johnson, K.F., Councill, E., Thorson, J.T., Brooks, E., Methot, R.D., Punt, A.E., 2016. Can autocorrelated recruitment be estimated using integrated assessment models and how does it affect population forecasts? Fish. Res. 183, 222–232.

Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M. A., Poos, J. J., Scott, F. and Scott, R. D. 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64(4), 640–646.

Kell, L.T., Kimoto, A. and Kitakado, T., 2016. Evaluation of the prediction skill of stock assessment using hindcasting. Fisheries research, 183, pp.119-127.

Knutsen, H., Jorde, P.E., Gonzalez, E.B., Eigaard, O.R., Pereyra, R.T., Sannæs, H., Dahl, M., André, C., Søvik, G. 2015. Does population genetic structure support present management regulations of the northern shrimp (Pandalus borealis) in Skagerrak and the North Sea? ICES Journal of Marine Science 72(3): 863–871.

Kolody, D. S., Eveson, J. P., Preece, A. L., Davies, C. R., & Hillary, R. M. (2019). Recruitment in tuna RFMO stock assessment and management: A review of current approaches and challenges. Fisheries Research, 217, 217-234.

Methot, R.D., Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142 (2013) 86–99.

Methot, R.D., Wetzel, C.R., Taylor, I.G., Doering, K.L., and Johnson, K.F. 2021. Stock Synthesis User Manual Version 3.30.17. NOAA Fisheries Seattle, WA, June 11, 2021

Mildenberger, T. K., Berg, C. W., Kokkalis, A., Hordyk, A. R., Wetzel, C., Jacobsen, N. S., Punt, A. E., et al. 2021. Implementing the precautionary approach into fisheries management: Biomass reference points and uncertainty buffers. Fish and Fisheries: 1–20.

Perälä T, Kuparinen A. 2017. Detection of Allee effects in marine fishes: analytical biases generated by data availability and model selection. doi/full/10.1098/rspb.2017.1284

Perälä, T., Huthchings, J. A. and Kuparinen, A. 2021. Allee effects and the Allee-effect zone in northwest Atlantic Cod. Biology Letters, 18 (2), doi.org/10.1098/rsbl.2021.0439.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A., Haddon, M., 2014. Management strategy evaluation: best practices. Fish and Fisheries, https://doi.org/10.1111/faf.12104

Siekmann, I., Sneyd, J., & Crampin, E. J. (2012). MCMC can detect non identifiable models. Biophysical journal, 103(11), 2275-2286.

Søvik, G. and Thangstad, T. 2021. Results of the Norwegian Bottom Trawl Survey for Northern Shrimp (Pandalus borealis) in Skagerrak and the Norwegian Deep (ICES Divisions 3.a and 4.a east) in 2021. NAFO SCR Doc. 21/001, Serial No. N7157. 38 pp. https://www.nafo.int/Portals/0/PDFs/sc/2021/scr21-001.pdf

Thorson, J. T. and Barnett, L. A. K. 2017. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. ICES Journal of Marine Science. doi: 10.1093/icesjms/fsw193

Thorson, J. T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research 210: 143-161.

Walters, C. J., Hilborn, R., & Christensen, V. (2008). Surplus production dynamics in declining and recovering fish populations. Canadian Journal of Fisheries and Aquatic Sciences, 65(11), 2536-2551.

4 pra.27.12 Barents Sea (ICES Subareas 1 and 2)

4.1 Stock ID and sub-stock structure

Northern shrimp (*Pandalus borealis*) in the Barents Sea and in the Svalbard fishery protection zone (ICES Subareas 1 and 2) is for management purposes considered as one stock. While the coastal and offshore parts of this stock is fished by different fleets, and new genetic studies in fact do indicate some separation between coastal shrimp and offshore shrimp, WKPAND was not in a position to recommend a change in the assessment or management unit at this time. However, it is strongly recommended that issues concerning stock structure be further investigated and considered in the future.

4.2 Issue list

Until 2006 management advice for this stock was formulated by qualitative assessment of trends in various indices of stock condition in response to the catch history and the predation by cod. An alternative quantitative assessment framework based on the work of Hvingel and Kingsley (2006) using a Bayesian version of a stock production model was introduced in 2006 and has been used since then. In addition to the NIPAG reports, detailed documentation of sampling, Input data and model runs over the historic assessment period, is found on the NAFO website in various SCR Doc's (Scientific Council Research Documents) Northwest Atlantic Fisheries Organization > Library > Science Council > SCR (nafo.int). The procedurals (documentation and otherwise) of this assessment have until now largely conformed to those of NAFO for historic reasons. This benchmark will, in addition to a full evaluation of input data and assessment models, also mark the transformation to the ICES practices regarding documentation (i.e., stock annex, TAF, etc.).

In advance of WKPAND 2022, an issues list for the assessment framework for this stock was developed by NIPAG: https://community.ices.dk/ExpertGroups/benchmarks/2022/WKPRAWN. This report is organized to address these issues in the order they appear in that document.

4.3 Ecosystem drivers

Population dynamics are subject to environmental, oceanographic and ecosystem processes. Data on processes relevant for a specific stock can, thus, potentially contain information that improves the estimation and prediction power within stock assessment. The two main ways of integrating such information is either as part of the standardization of stock indices or directly as an input time series to the stock assessment model.

Temperature and cod (*Gadus morhua*) predation have been suggested as the two main drivers of shrimp stock dynamics (Hvingel 2006a). This was first investigated for the Barents Sea stock by Aschan *et al.* (2006) and Hvingel (2006b) however, without finding significant correlations with the dynamics of the shrimp stock. For the WKPAND 2022 new and more extensive analyses were performed to test for relationships between i) shrimp density, cod density and bottom temperature at the same location as observed during the Barents Sea Ecosystem Survey (BESS); ii) changes in estimated shrimp biomass and cod biomass as well as temperature, and other environmental indices.

The data exploration showed little to no link between shrimp observations and cod density (Figure 4.3.1) or bottom temperature (Figure 4.3.2). This was confirmed in the statistical modelling of the survey index, where both cod and temperature were tested during model selection as covariates but found to be not relevant (see section 2.4.2.1).

No clear signals were found in the analysis of shrimp biomass trends in relation to cod biomass and environmental indices. The slight negative trend between cod SSB and shrimp (Figure 4.3.3) was found to be only marginally significant, and it was determined that the potential signal in the data is primarily driven by the huge peak in cod SSB during a period of slightly decreased shrimp biomass (Figure 4.3.4). Over the entire time series, there is little indication of a clear inverse relationship that would justify the inclusion of cod as predictor of shrimp in the assessment model. Similarly, no statistically relevant relationship between environmental indices and shrimp biomass was found for the same year (Figure 4.3.5) or lagged with up to 3 years (not shown).

Based on these analyses, we concluded that the link between existing ecosystem and environmental data and shrimp is insufficient to justify inclusion in either the standardization of the survey index nor in the assessment model. This may contradict knowledge about mechanistic links between shrimp dynamics and its environment, especially in the form of cod predation. However, finding clear signals for predation or environment as driver of stock dynamics in available data has been often challenging. This indicates, most likely, not so much that such links do not exist, but rather that the available data is inadequate in terms of spatial and temporal resolution, quality, or time series length to detect statistically significant patterns within complex processes that are subject to a multitude of drivers. Generally, total biomass (both as observed density or estimated stock biomass) may be too coarse to determine a clear link, compared to more specific metrics such as recruitment/cohort strength. It was therefore concluded that i) *there is currently no sufficient basis to include available ecosystem or environmental information into the assessment*; ii) *the issue should be re-evaluated once new information becomes available*.

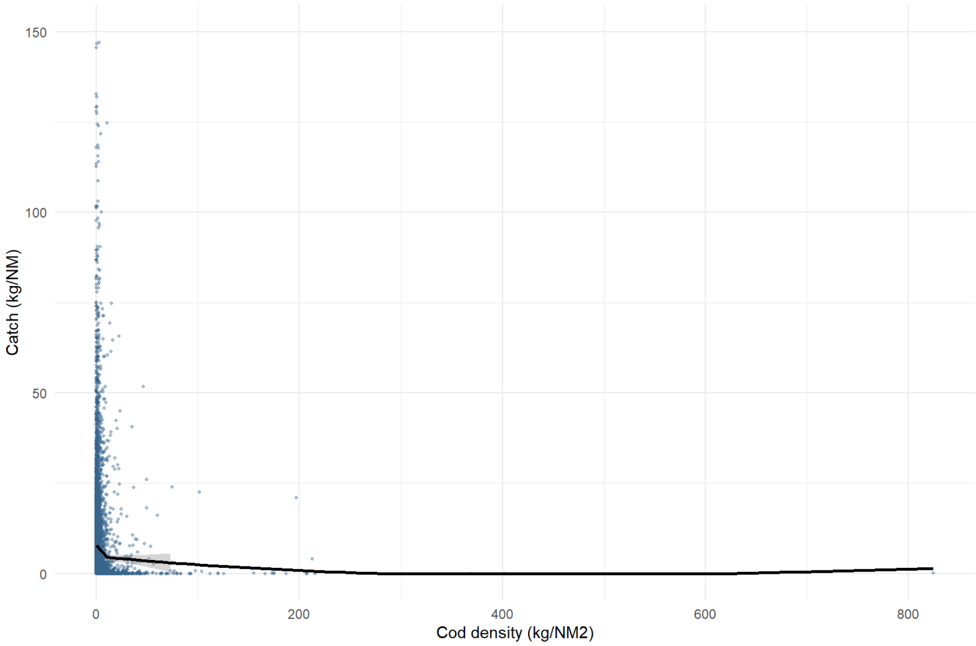


Figure 4.3.1. Relationship between cod density and shrimp catches observed during the BESS. The blue dots represent each the combined registration of cod and shrimp at one survey station during the period 2004 -2021, the black line is an overlaid trend line based on a GAM smoother spline.

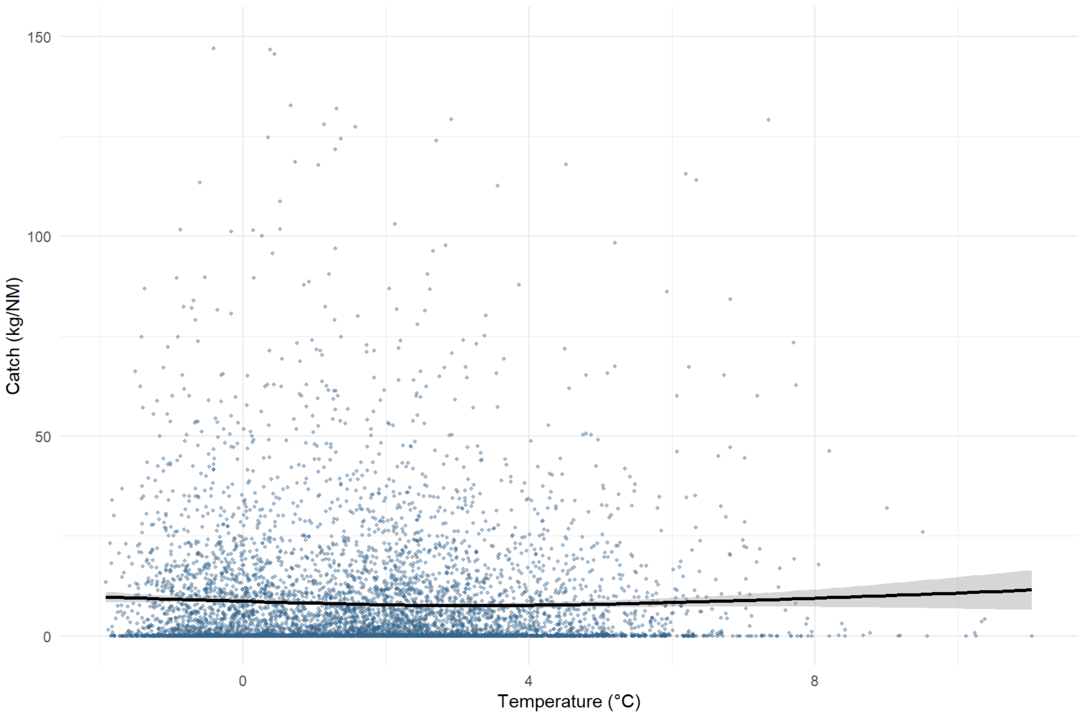


Figure 4.3.2. Relationship between bottom temperature and shrimp catches observed during the BESS. The blue dots represent the combination of shrimp observations at one survey station during the period 2004–2021 and modelled mean monthly bottom temperature at the same location, the black line is an overlaid trend line based on a GAM smoother spline.

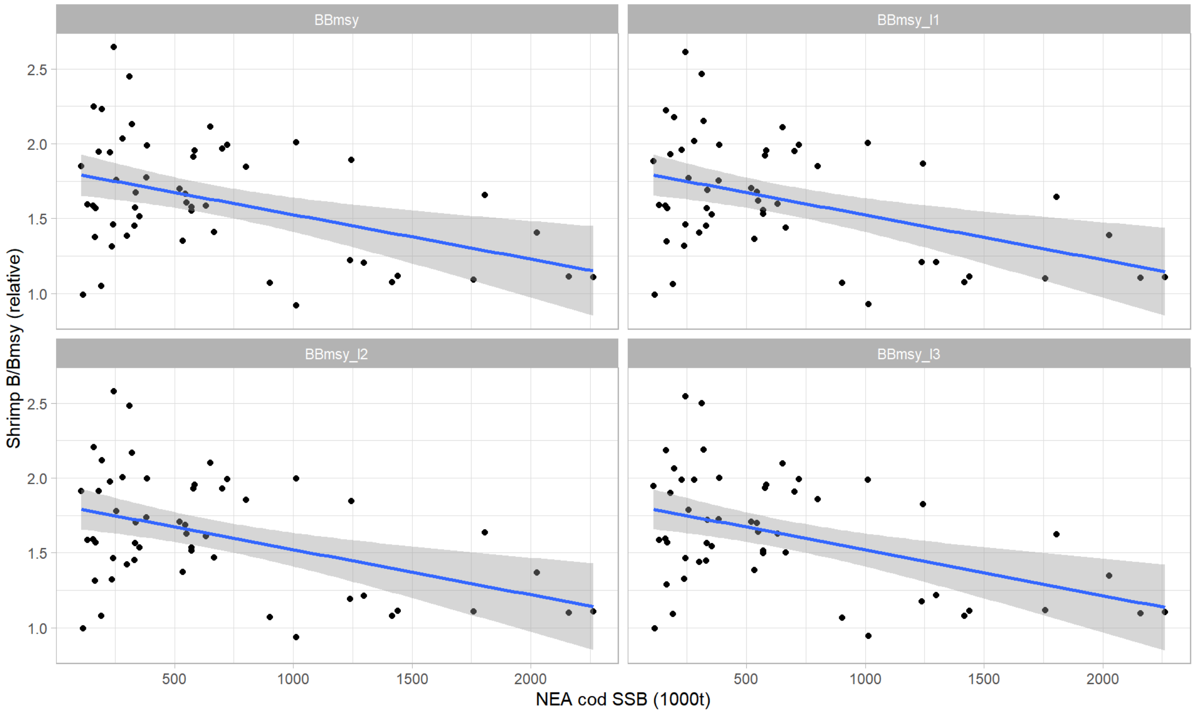


Figure 4.3.3. Relationship between estimated NEA cod spawning stock biomass and shrimp biomass in the same year or lagged by 1, 2 or 3 years. Each dot represents cod and shrimp biomass in a given year, while the blue lines and shaded area show linear trend lines and 95% confidence intervals, respectively.

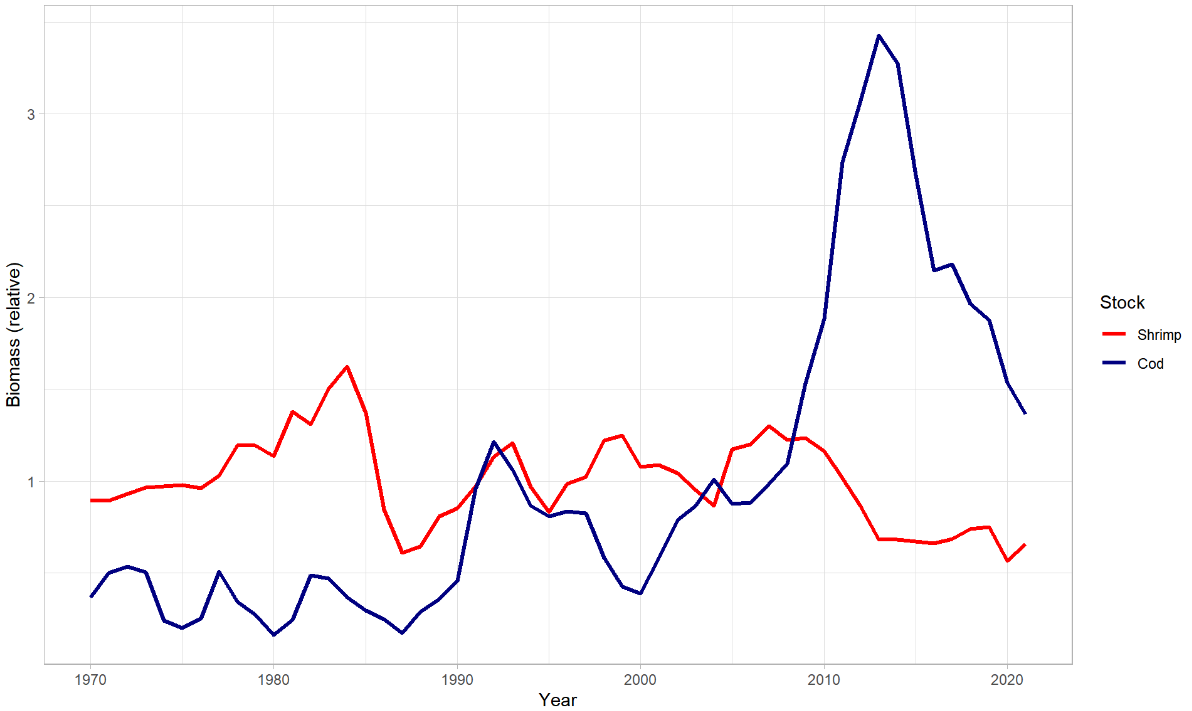


Figure 4.3.4. Stock biomasses of NEA cod and shrimp for the period 1970 to 2021. Biomasses were standardized to their respective means.

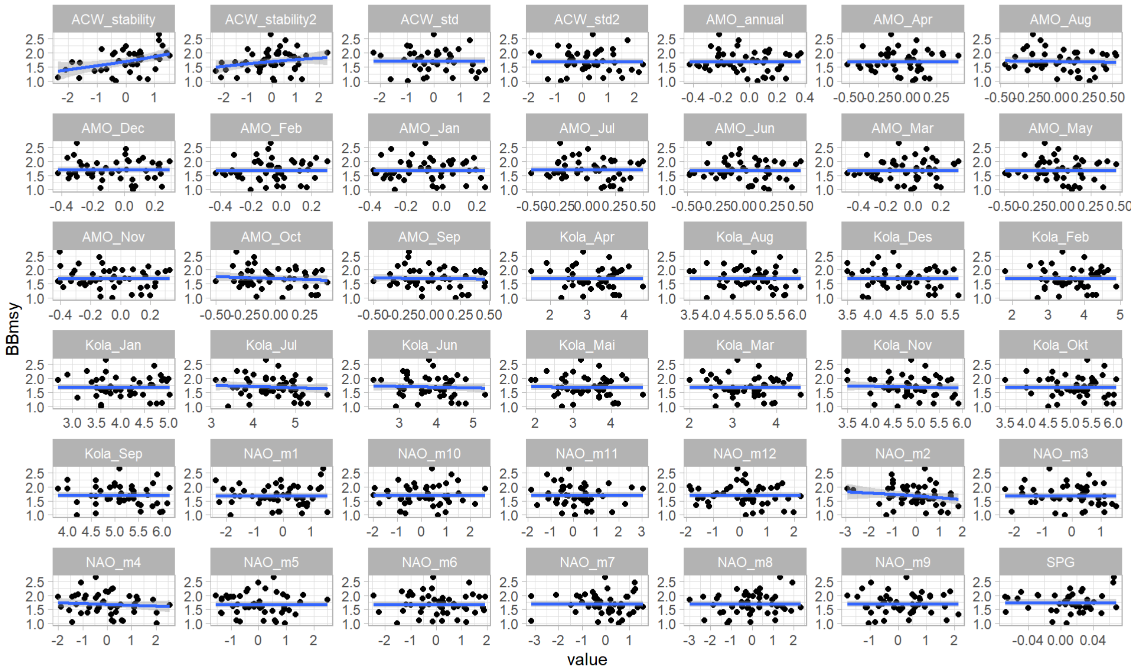


Figure 4.3.5. Relationship between environmental indices and shrimp biomass in the same year. Each dot represents environmental index and shrimp biomass in a given year, while the blue lines and shaded area show linear trend lines and 95% confidence intervals, respectively. Included environmental indices are: Norwegian coastal current (ACW, stability and strength), Atlantic Multidecadal Oscillation (AMO, annual and monthly), Kola temperature index (Kola, monthly), North Atlantic Oscillation (NAO, monthly), Subpolar Gyre Index (SPG).

4.4 Stock Assessment

4.4.1 Catch – quality, misreporting, discards

The fishery is multinational: Norwegian and Russian vessels exploit the stock in the entire area while vessels from other nations are restricted to the Svalbard fisheries zone and the loophole (international waters). Catches have ranged between 19 000 and 128 000 tonnes since the mid-1970s. Historically Norway has accounted for the major part (~75-95%) of the landings, however, in the recent years the Norwegian proportion has decreased to less than 40% while EU, Russia, Greenland and Iceland share the rest.

The fishery is regulated by TACs in the Russian zone and effort control in the Norwegian and Svalbard zones: licenses are required for the Russian and Norwegian vessels and the fleets operating in the Svalbard zone are regulated by number of effective fishing days and number of vessels by country. Minimum cod-end mesh size is 35 mm. Other species and small shrimp are protected by mandatory sorting grids (19 mm bar space) and by the temporary closing of areas with excessive by-catch of juvenile cod, haddock, Greenland halibut, redfish, and shrimp <15 mm carapace length (measured in catch samples taken by independent observers). In the international waters of the 'Loophole', NEAFC regulations require a minimum mesh size of 40 mm, a cod end of 44 mm and an obligatory sorting grid bar space of 22 mm. In this area there is no effective limit on the overall level of fishing effort or an overall quota.

The official catch statistics provides estimates of overall landings. Discard of shrimp cannot be quantified but is believed to be small as the fishery is not limited by quotas. There is little incentive for misreporting. The fishery being without major discarding problems or variable misreporting, reported landings are considered equal to actual catch and is entered into the assessment model as error-free.

4.5 Surveys

4.5.1 Barents Sea ecosystem survey

The Barents Sea ecosystem survey (BESS) is a joint Norwegian-Russian ecosystem survey conducted since 2004 and covers the entire Barents Sea with bottom trawl station grid. A BESS total biomass index that has been used in the previous assessments of Barents Sea shrimp was based on a design-based swept-area estimate. This approach is widely used and assumes that observed density in a sub-stratum is representative for the entire area it is raised to. The main advantage of this method is that it in principle constitutes and unbiased estimate, and it requires only straight-forward calculation of descriptive statistics. Potential downsides are the need for careful stratification to avoid biases and the lack of analytical statistics that can utilize information from covariates.

The key issue with the current method in the case of Barents Sea shrimp, however, has been incomplete survey coverage in several years due to technical problems. Design-based survey indices tend to perform poorly when incomplete coverage results in under-sampled strata with too few or no observations at all. Typically, the latter problem has been addressed by using ad-hoc corrections, inflating observed biomass with past biomass proportions to compensate for the missing areas.

The use of (geo-)statistical methods to estimate survey indices has gained traction in recent years, to a large degree following the development of efficient methods to apply mixed-effect models that include spatio-temporal correlation, notably in R-based applications such as R-INLA (Beguin et al. 2012), sdmTMB (Anderson 2019) and VAST (Thorson 2019). These approaches allow to account for spatial- or spatio-temporal variation in data through Gaussian Markov fields, in addition to modelling the observed densities as a function of other (linear or additive) fixed and random effects. Consequently, modelled standardized indices have been shown to reduce error and improve predictive power, especially when survey coverage is incomplete (Breivik *et al.*, 2021).

A statistical approach as alternative to the current design-based method to estimate standardized survey indices was explored. In an initial step, an GLMM without the inclusion of spatio-temporal correlation was tested, modelling shrimp density with a compound Poisson-gamma tweedie distribution in glmmTMB (Brooks et al. 2017). Bottom depth and time of day were selected as explanatory variables, in addition to strata as random intercepts. The model performed well, however clear indications for spatial correlation were found in visual inspection of residuals and the sample variogram of residuals (Figure 4.5.1). This confirmed an approach that can model spatio-temporal correlation is required to fully capture the spatial variation in the survey data. The packages R-INLA, sdmTMB and VAST were tested for this purpose, and after promising results sdmTMB was selected for further analysis as the most suitable and user-friendly option for this specific case.

The selected sdmTMB model was a GAMM with non-linear splines for bottom depth and time of day, and spatial random fields by year linked through an AR1 process:

$$Density_{i,s} = Year_i + s(Depth_{i,s}) + s(Daytime_{i,s}) + w_{i,s}$$

at year i and location s, with a $Year$ intercept (categorical fixed effect), the natural splines for continuous fixed effects D$epth$ and $Daytime$ restricted to three degrees of freedom, and spatially correlated random effects with mean 0 and a Matérn covariance matrix Σ_k by year $w_{i,s} = f(x_i, y_i \sim \text{AR1}) \sim GF(0, \Sigma)$.

The resulting index showed, overall, similar interannual changes as the previously used design-based index (Figure), although with some deviations. In addition, the results showed a clear link between bottom depth and shrimp density, and a relatively minor effect of time of day on shrimp catches. The distribution of shrimp over time was very consistent (Figure) with little change in the center of gravity (Figure), reflected in an estimated AR1 coefficient of 0.67. The robustness of the method was tested with retrospective analysis, i.e., by peeling off 1 to 5 years from the survey data and re-estimating the index from the shortened datasets. The result showed that the modelled survey index is robust, both in terms of estimation consistency (Figure) and convergence. *It was therefore decided to replace the design-based biomass estimate with the sdmTMB modelled survey index in the stock assessment.*

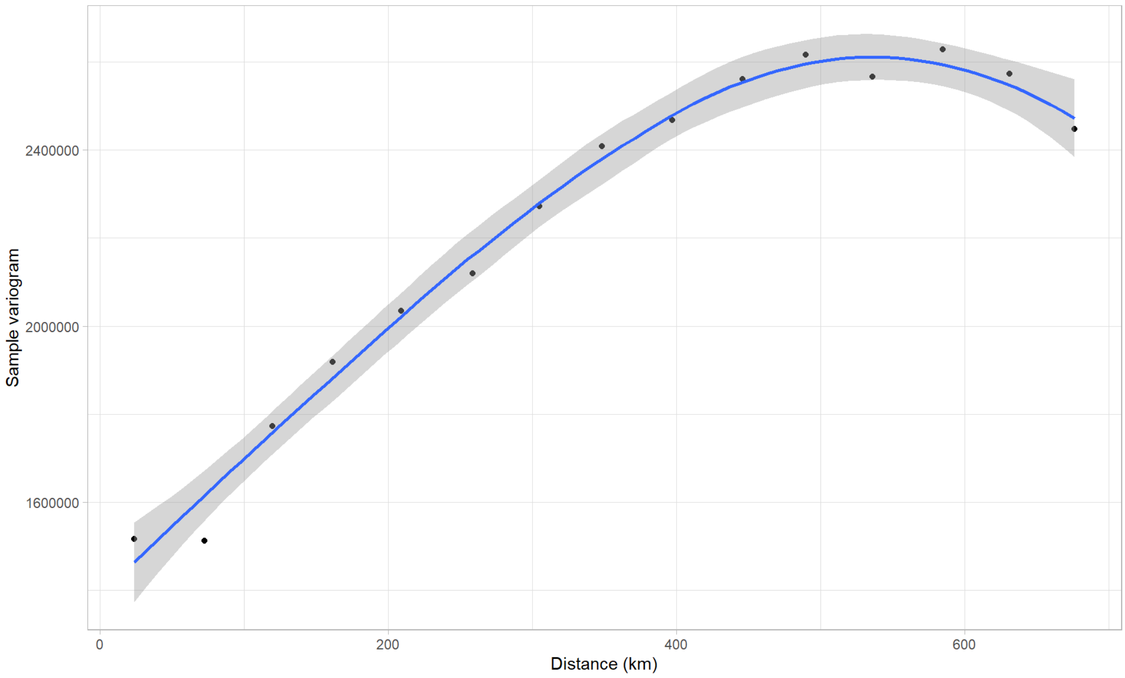


Figure 4.5.1. Sample variogram of residuals of GLMM to BESS shrimp density data.

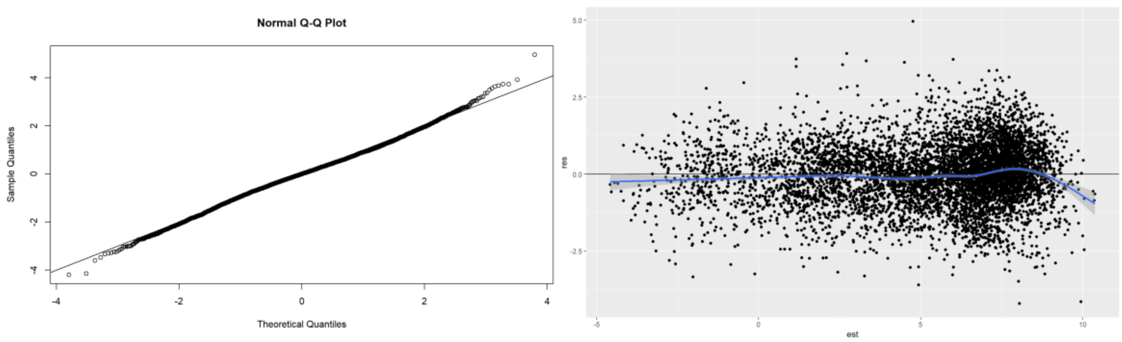


Figure 4.5.2. Residual distribution of the fitted sdmTMB models, using simulated residuals.

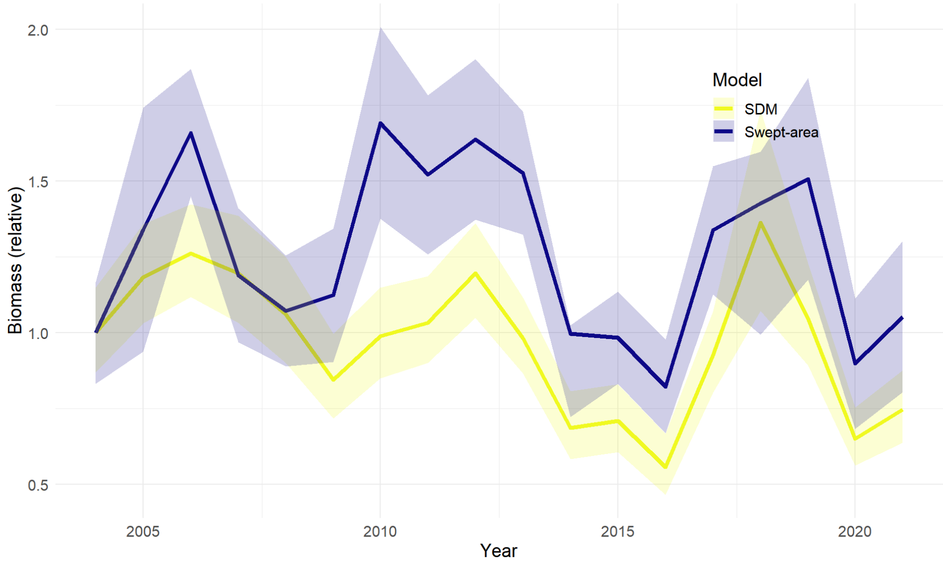


Figure 4.5.3. Biomass indices of BESS shrimp data standardized with the current design-based approach (blue) and a statistical approach using sdmTMB (yellow). Both indices are shown relative to their value in 2004. Lines indicate mean estimates, shaded area the corresponding 95% confidence intervals.

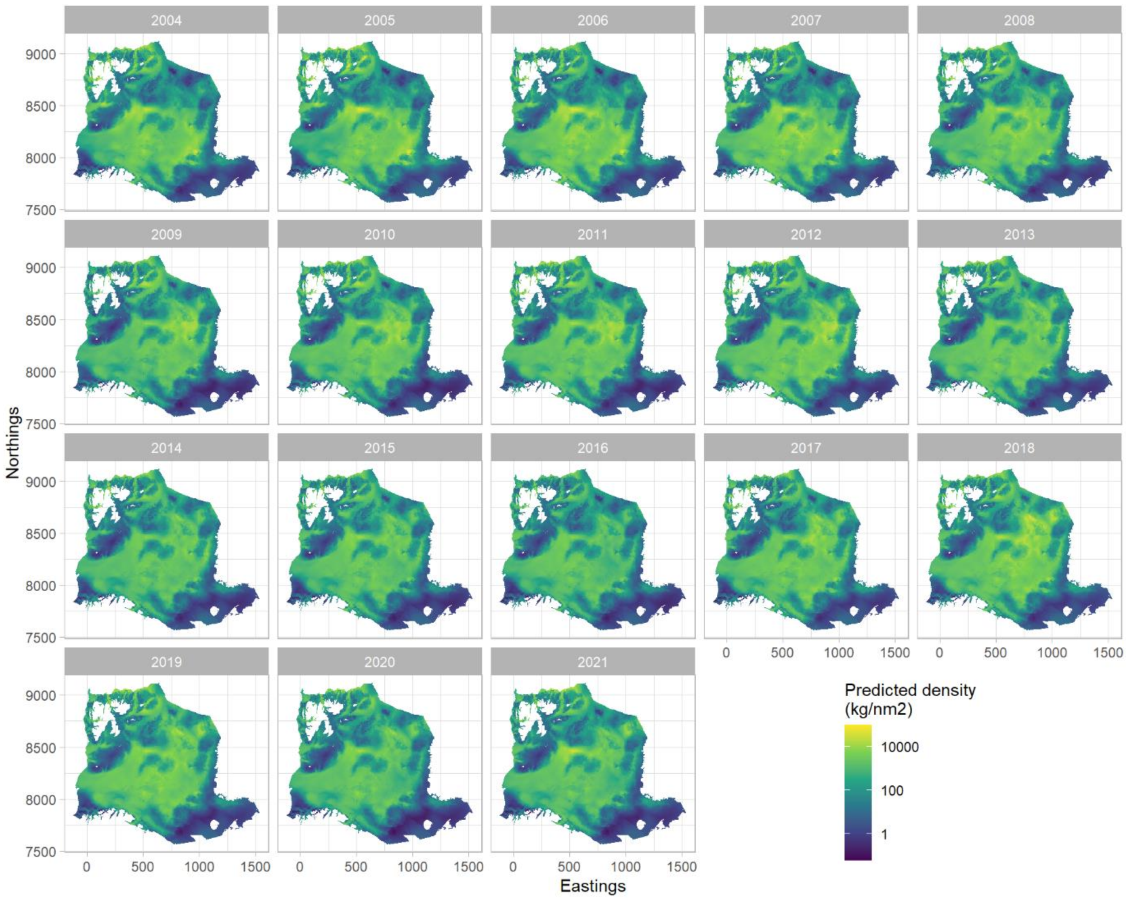


Figure 4.5.4. Density of shrimp as predicted by the sdmTMB model fitted to BESS data.

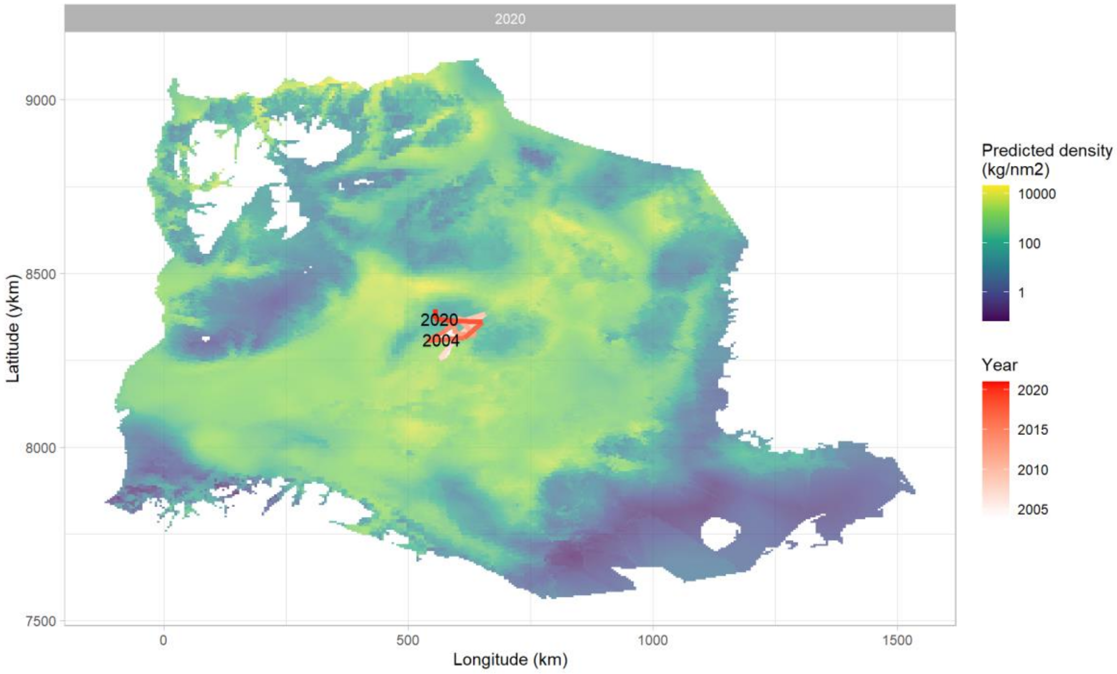


Figure 4.5.5. Predicted density of shrimp in the Barents Sea and center of gravity for the years 2004 – 2021, estimated with sdmTMB model fitted to BESS data.

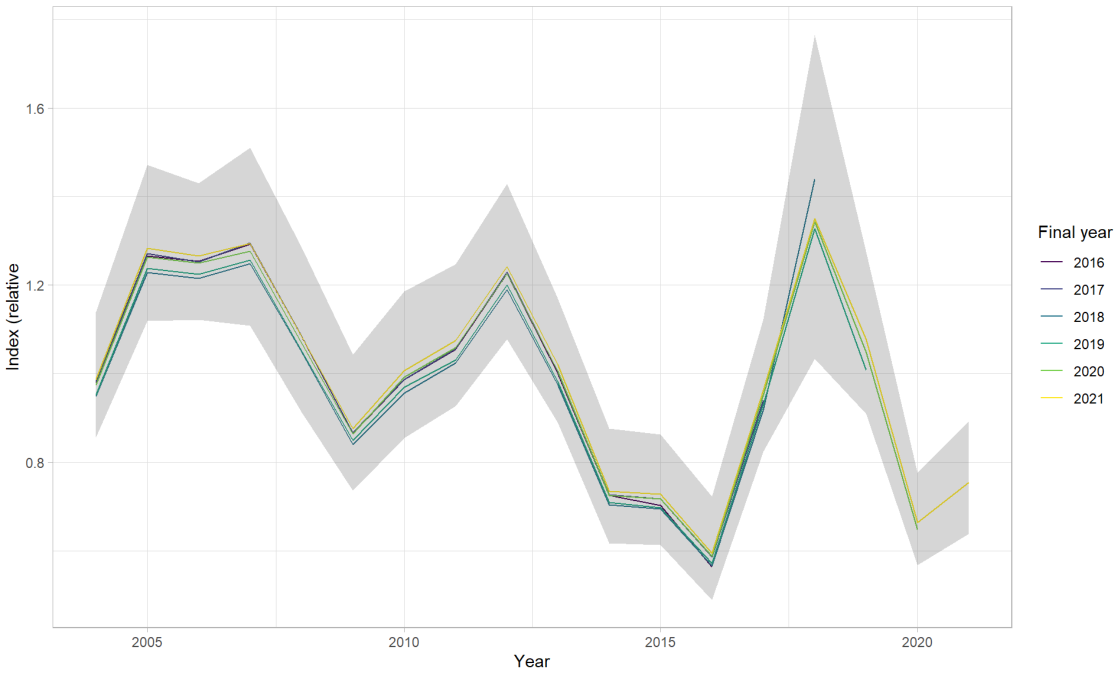


Figure 4.5.6. Retrospective analysis of the survey index modelled with sdmTMB, re-estimating the index from BESS data with 2016 – 2021 as final years. 2021 corresponds to the base case shown in Figure .

4.5.2 Barents Sea demersal fish survey in winter

The demersal fish survey in winter (WS) is a bottom trawl survey conducted annually in the Barents Sea between January and March. The survey covers the accessible areas during the winter, i.e. the parts of the Barents Sea not covered by sea ice. Thus, survey coverage tends to vary and is mainly restricted to the southern and western parts of the Barents Sea (Figure). Number of stations and collaborative efforts between Norway and Russia have increased since the 90ies, mirroring the survey design of the BESS for the accessible areas. Accordingly, there are relevant observations of shrimp within a substantial part of the stock distribution, making the WS an additional source of fisheries-independent data.

Data exploration showed that shrimp observations on the WS are comparable in total number and proportion of all stations to the ones on the BESS (**Error! Reference source not found.**). Specifically, there is substantial overlap in the coverage and registered shrimp catches in stratum 2 that covers a large area in the southwestern Barents Sea (Figure).

A comparison of swept-area total biomass estimates based on data from the BESS and WS showed significant discrepancies across the entire survey area (**Error! Reference source not found.**), despite a similar biomass level and an improved correspondence in the most recent years. The deviations can be explained due to the incomplete coverage of the WS. However, when focussing on stratum 2 where the coverage of the WS is most consistent, a strong within-year correlation between the two survey indices emerged (Figure 6).

Based on the data exploration, it was concluded that the WS data is insufficient to provide a survey index that is representative for the entire stock area. However, the WS may be able to provide signals of stock development early in the year that may increase the overall robustness of survey information and provide important information for the assessment year if BESS data is missing or incomplete when the stock assessment is conducted. The limited time available between the end of the BESS and the NIPAG assessment meeting, as well as force majeure issues during the BESS, have been a challenge in the past, resulting in increased uncertainty in the survey index. Combining the WS and BESS survey data in a joint survey index was therefore proposed to i) make use of the WS data while circumventing the coverage limitations that would be unavoidable in a separate WS survey index; ii) strengthen the current survey index by adding additional data that can provide information on stock development in the current assessment year, especially when BESS data is limited or missing for the stock assessment. The latter may also increase the flexibility in the possible timing of the assessment, which has been so far constrained by the timing of the BESS.

The inclusion of WS data into a joint BESS/WS survey index was tested as part of the re-evaluation of the index standardisation methods. The same sdmTMB model configuration as for the BESS survey index was used, except that the model was fitted to shrimp density observation in both datasets combined into one, and a random intercept was added account for potential differences in the survey time series. The resulting index is very similar to the one estimated from BESS data alone and in addition consistent back in time (Figure). However, the retrospective analysis and tests with altered configuration of the spatial random fields revealed some potential convergence issues, likely due to the substantially increased data quantity. *The conclusion was therefore to use the modelled BESS index as the baseline input in the stock assessment model while also testing the effects of using the combined index instead.*

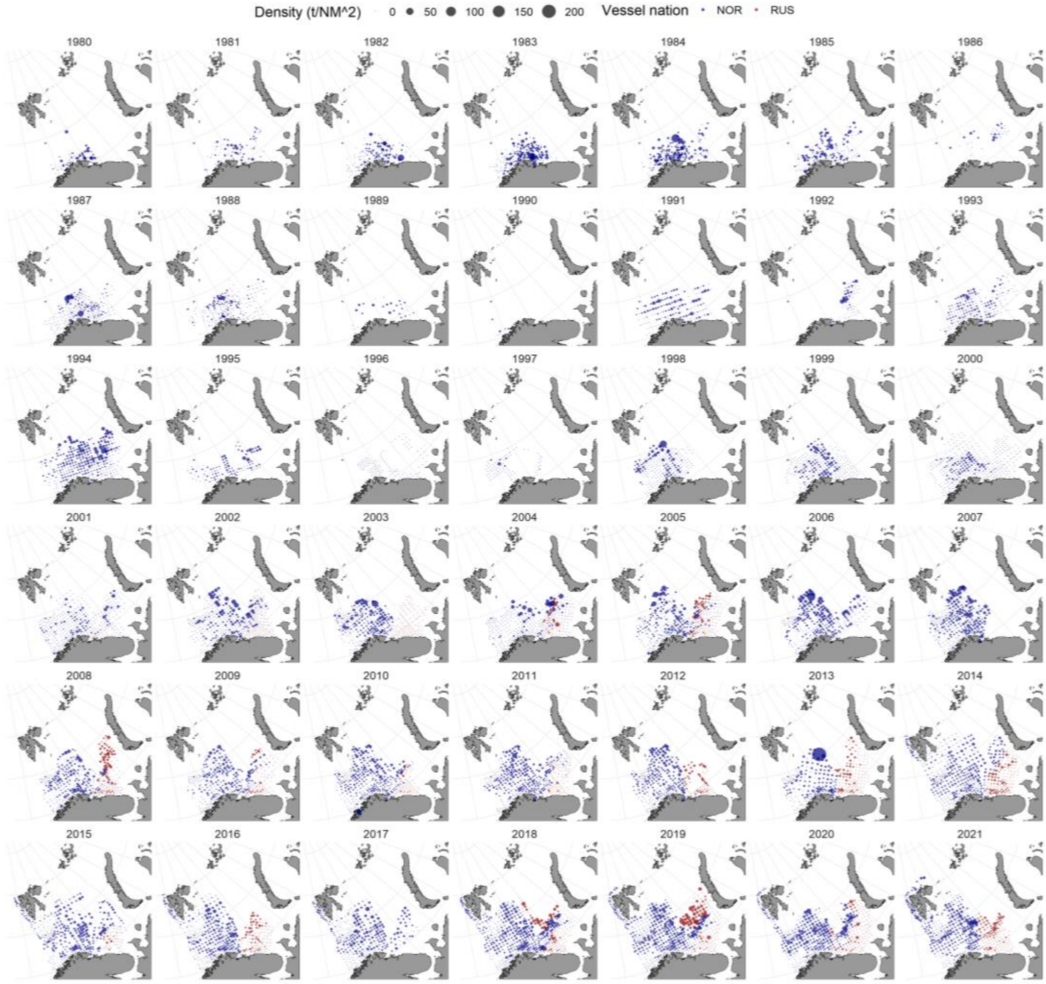


Figure 4.5.7. Bottom trawl stations at the demersal fish survey in winter in the Barents Sea, scaled to density of the registered shrimps. Colours indicate Norwegian and Russian vessels.

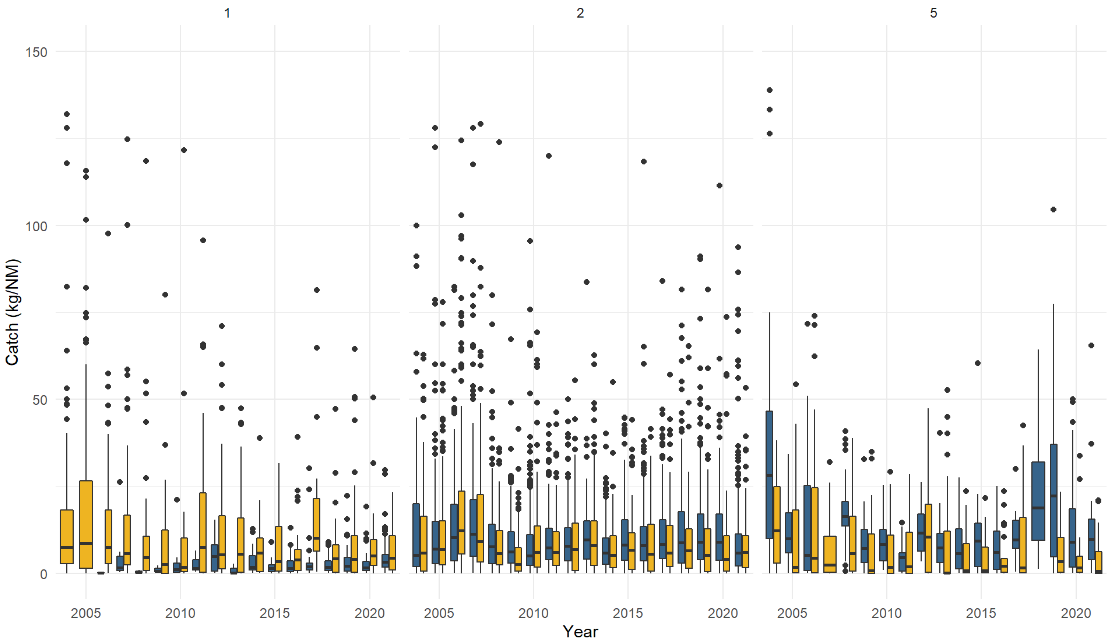


Figure 4.5.8. Boxplots of standardised shrimp catches on the BESS (yellow) and WS (blue) per year, separated by stratum. Shown are strata 1, 2 and 5, while 3 and 4 were excluded due to very limited coverage of the WS.

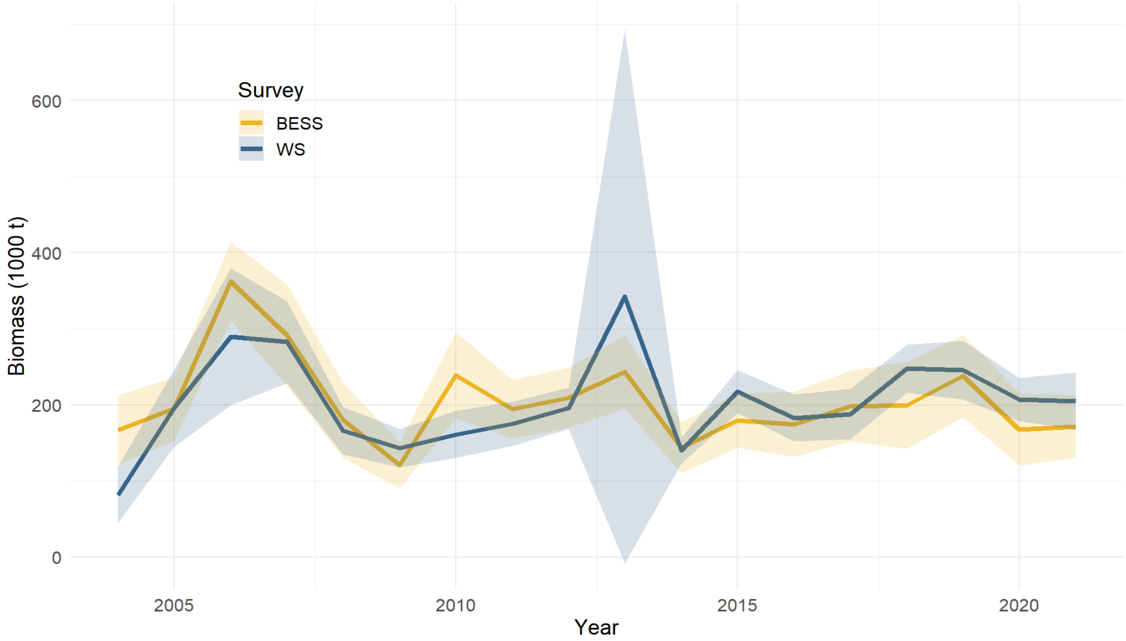


Figure 6. Swept-area biomass estimates in stratum 2 based on the data from the BESS (yellow) and WS (blue).

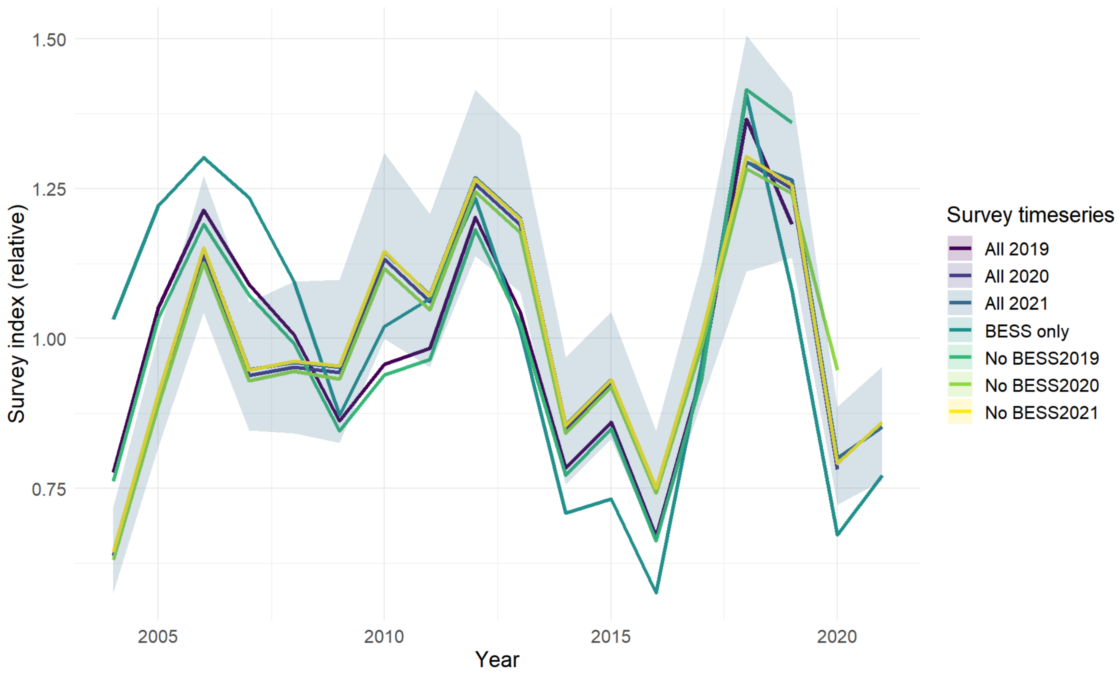


Figure 4.5.10. Biomass indices of Barents Sea shrimp standardized with a statistical approach including spatio-temporal correlation using sdmTMB (yellow), and combined data from the BESS and WS. The indices were estimated from the entire BESS and WS datasets from 2004 to 2019, 2020 and 2021, including ("All") or excluding ("No BESS") the BESS data in the final year. In addition, the index using BESS data only is shown (equivalent to Figure Error! Reference source not found.**), Both indices are shown relative to their value in 2004. Lines indicate mean estimates, shaded area the corresponding 95% confidence intervals (only shown for the bvase case "All 2021").**

4.6 Commercial CPUE standardization

The assessment has included a CPUE index based on a catch and effort time series from the Norwegian fleet starting in 1980. The fishery has undergone some substantial changes over the 40 years of the time series, notably a contraction in terms of participating vessels and fishing area during the second half of the time series. In recent stock assessments (NAFO/ICES 2020) some odd patterns have been observed for the estimated CPUE in the assessment year, likely linked to data filtering and the limited sample size due to few vessels active in the fishery. In addition, the collection of logbook data has been reformed in Norway in 2011 through the introduction of electronic logbooks that contain higher resolution data (by haul instead of day). This resulted also in changed gear codes, which have been included as categorical variables in the CPUE standardization, specifically that they do contain information on number of gears. Instead, number of gears have been registered separately from 2011 onward and this information was thus not adequately accounted for in the index standardization. Due to these changes in the fishery and data, a re-evaluation of the CPUE index standardization was added to the issue list, focusing on the modelling approach, the data filtering and spatial structure, and whether the CPUE index accounts sufficiently for technological creep.

The CPUE index standardization was previously conducted with a GLM that included in addition to a year effect the following other categorical variables: gear code, month, stratum (based on the old Norwegian shrimp survey) and vessel ID. Because modelling spatial and ID clusters as random effects has become the more suitable and accepted model structure, mixed-effects models were selected as modelling approach for the revised index, using glmmTMB. Comparative analysis showed that the effect of this change on the estimated index was negligible but resulted in a general improvement of model fit. Modelling CPUE as a Gamma distribution compared to the current log-normal distribution was tested but found to not fit the data better. A model selection was conducted to determine the best combination of explanatory variables, testing month (categorical or continuous) or Julian day to account for seasonal effects, and vessel size or engine power, in addition to the categorical year effect. Gear codes were replaced with explicit gear numbers (categorical) when including electronic logbook data.

All steps of data exploration and CPUE standardization were conducted with the CPUE data in the old logbook format (as used in the 2021 assessment), the electronic logbook data alone (2011-2021), and a combined dataset (old format for 1980–2010, with gear codes translated to gear numbers; electronic logbook format for 2011–2021). The selected models and resulting indices showed a good alignment (Figure), confirming that substituting the old logbook format with the new electronic logbook format does not per se alter the CPUE index (as can be expected given that the underlying data are the same). However, the CPUE index estimated from the combined dataset reveals a slightly stronger decrease in CPUE in the most recent period, likely because the number of trawls were not adequately accounted for in the old format after 2010, resulting in an overestimation of CPUE that has been corrected with the revised index estimated from new format. The underlying model was specified as:

$$\log(CPUE_{i,s,v,h}) = Year_i + Gears_{i,s,h,v} + Size_{i,v} + Month_i + (1|Vessel_v) + (1|Stratum_s)$$

with CPUE in year i, in stratum s, for vessel v, and haul h modelled with a categorical year effect, the number of gears (categorical) and size of vessel, the month of the year (categorical), and normally distributed random intercepts for vessel ID and stratum. The CPUE index was predicted as the population level mean per year for month 6.

The distribution over time of survey and CPUE data and the sensitivity of survey and CPUE indices to strata definitions were investigated, comparing survey strata systems and statistical grid cells as random intercepts to account for spatial clustering. Indices were robust to strata definition despite some relevant shifts in distribution over time in the fishery-dependent data. It was decided to use the strata definition from the BESS in the CPUE. Sensitivity analysis of the CPUE standardization was conducted to determine the robustness of the CPUE index to filtering and modelling specification. Data filtering has been used to ensure that only representative vessels are included in the CPUE index, which is relevant due to a relevant proportion of opportunistic vessels that only occasionally fish for shrimp. The index was found to be insensitive to the relevant range of filtering thresholds, showing that the default settings (at least 3 years present in the fishery with at least 20 observations per year) are robust (Figure).

The model aims to correct for technological creep through the inclusion of number of gears, vessel size and vessel ID. Additional unaccounted technological creep e.g., through improvement of fishing power of individual vessels during the period active in the fishery, was explored by inspection of vessel-specific residuals, focusing on vessels with the longest participation in the fishery. No trends or problematic patterns were found. However, indications for some minor spatial correlations were found in the distribution of residuals (Figure) and the corresponding semi-variogram. A GAMM in sdmTMB that includes spatio-temporal correlation was tested but not implemented as the final model due to convergence issues, likely due to the large amount of data (N > 300,000) and the limited spatial resolution for all data before 2011 (only midpoints of statistical grid cells). I*t was therefore decided to use the GAMM implemented in glmmTMB as the new revised CPUE index but keep exploring possibilities to account for the remaining spatial correlation in the future.*

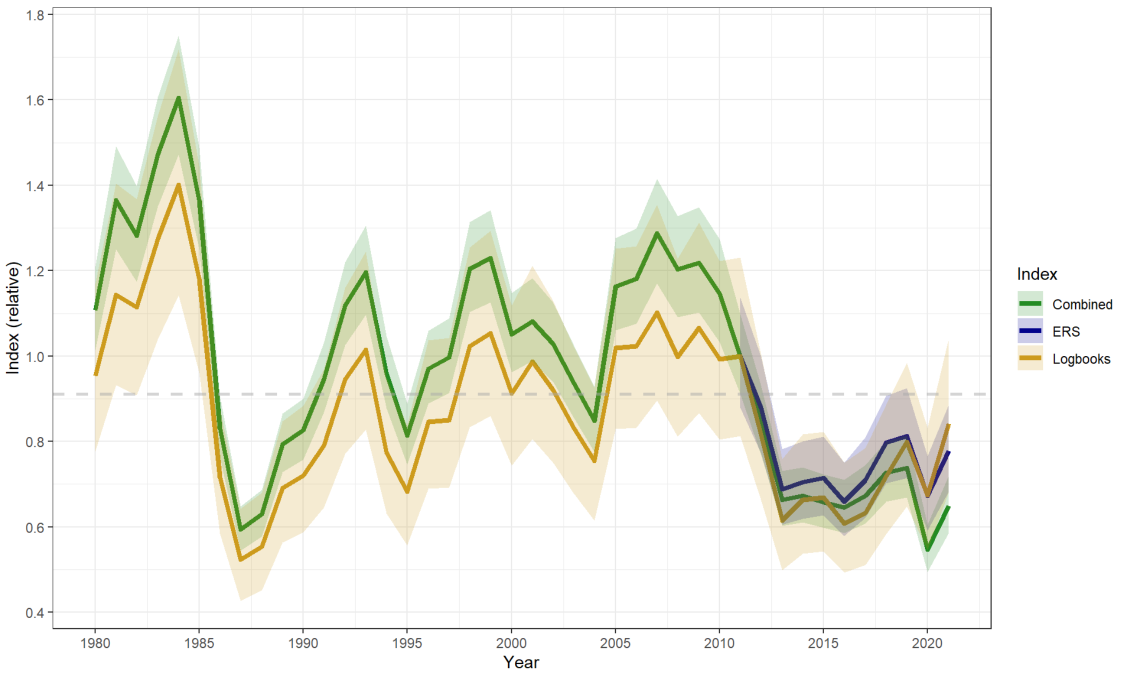


Figure 4.6.1. Standardized indices estimated from selected GLMMs fitted to CPUE data in the old logbook format (yellow), the electronic logbooks only (blue) and a combined dataset (green).

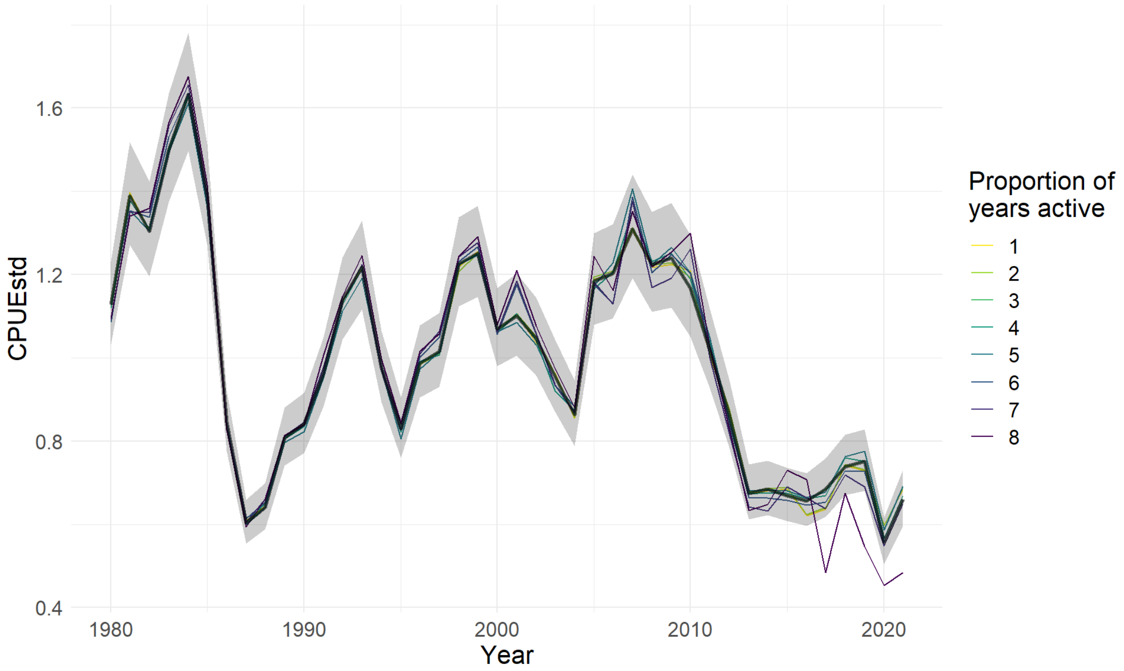


Figure 4.6.2. Sensitivity of CPUE index estimate to filtering of data in terms of the minimum period a vessel is active. Default (black line and shaded area) is 3 years.

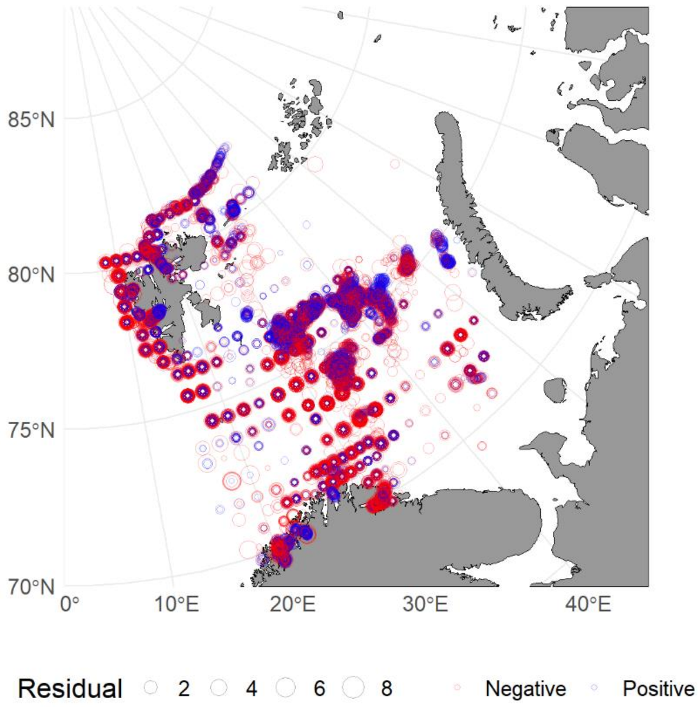


Figure 4.6.3. Spatial distribution of residuals of selected GLMM, showing the distribution of negative (red) and positive (residuals) scaled to their absolute value.

4.7 Model priors

The Bayesian implementation of the assessment model (OpenBugs) has full flexibility in the choice of priors, while the frequentist implementation (SPiCT) is restricted to lognormal priors. The considerations below may therefore be applied differently depending on the implementation used.

4.7.1 Carrying capacity (K)

An Informative prior for carrying capacity, K, was constructed based on a K-estimate for the West Greenland shrimp stock. The West Greenland shrimp are living under similar environmental conditions – e.g. sea temperature regimes which are considered important drivers for growth rates and life histories (Shumway *et al.,* 1985). However, the level of predators (especially cod) at west Greenland has in recent years been very low and the K-estimates have increased in the annual updated assessment model runs – likely as a consequence (c.f. the annually updated NAFO SCR Docs addressing the assessment of the West Greenland shrimp stock found at Northwest Atlantic Fisheries Organization > Library > Science Council > SCR (nafo.int)). An estimate of K from a time period (until 2002) more comparable to the Barents Sea with varying but relatively large cod stock, is available in Hvingel and Kingsley (2006). This estimate is therefore used together with estimates of habitat size and relative habitat quality as the basis for the K prior.

Habitat size and quality defines the total biomass of shrimp that may exist at an unfished equilibrium state i.e., at carrying capacity. Off West Greenland, depths between 150 and 600 m are considered as suitable shrimp habitat. The habitat area, equal to the area surveyed annually by the West Greenland shrimp survey (WGS), is estimated to 136 000 km^2. The corresponding size of the shrimp habitat in the Barents Sea may be estimated as the fraction of the total survey area (1.5×10^6 Km^2) where shrimp is detected. Modelled distribution area by year based on presence/absence data from BESS shows presence of shrimp is around 66% of that area (Figure 4.7.1) equal to 993 000 km^2.

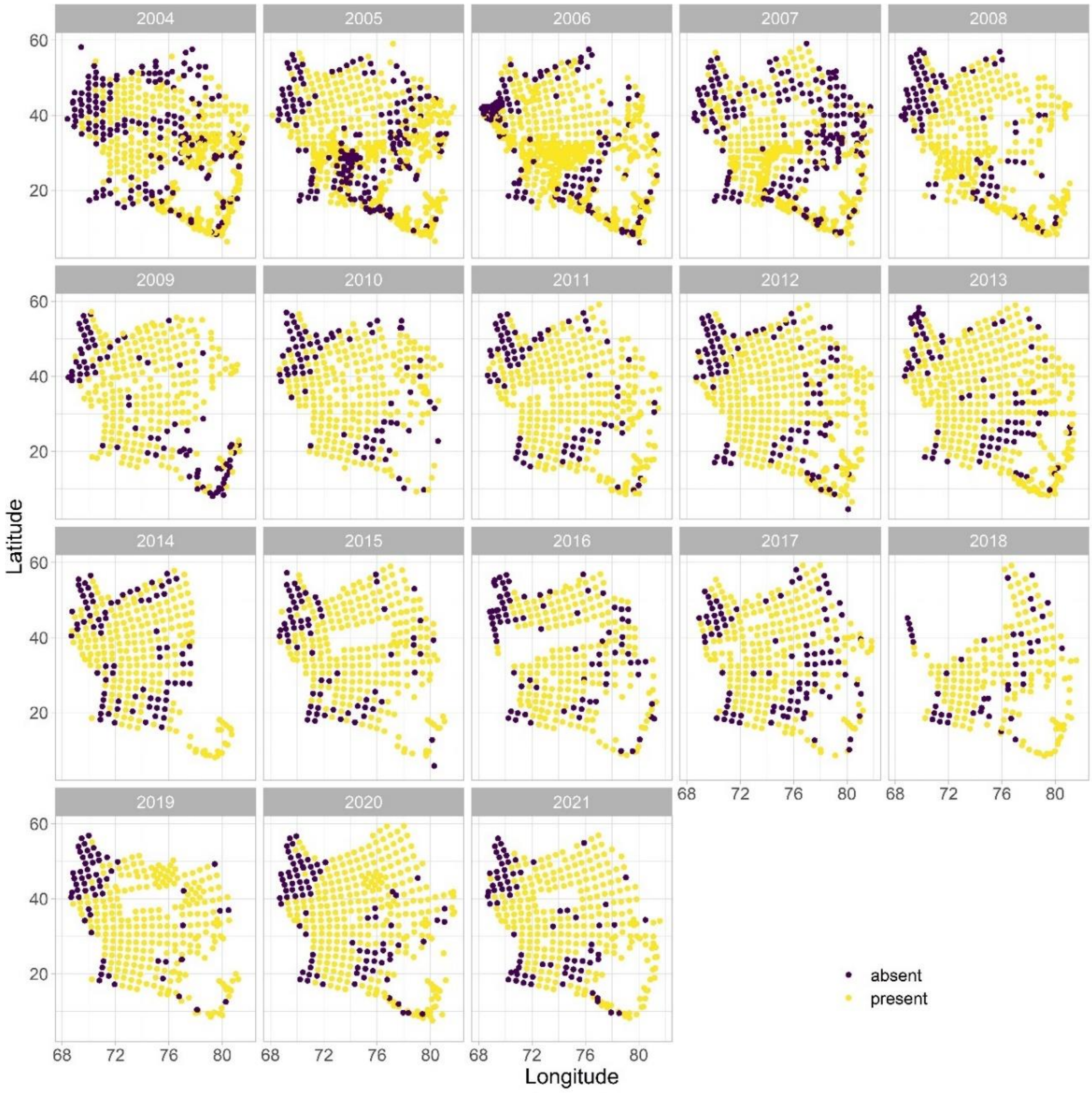


Figure 4.7.1. Presence/absence of shrimp in trawl hauls taken in the BESS survey 2004 to 2021.

The density measured in the WGS (Burmeister and Riget, 2020) over the period 1988 to 2002 range between 1.1 and 3.1 tonnes shrimp/km^2 with an average of 1.784 t/km^2. In the Barents Sea the mean density recorded in the BESS series is 0.547 t/km^2.

The two surveys, WGS and BESS, are conducted in a similar fashion using a "shrimp trawl". The densities of shrimp are calculated based on wingspreads and distance trawled (trawl on the bottom) i.e. area swept. Assuming that the catchabilities of two surveys by area swept are similar, the relative habitat quality of the Barents Sea is on average 0.547 / 1.784 = 0.31 times that of West Greenland.

The effective shrimp habitat area of the Barents Sea (distribution area times relative habitat quality) is thus 0.31 times 993 000 km^2 equal to 308 000 km^2.

The estimated posterior for K for the West Greenland shrimp stock (Hvingel and Kingsley, 2006) had a median of 728 ktonnes and 95% of the distribution between 300 and 2500 ktonnes. An informative prior for K for the Barents Sea stock may then be derived by multiplying the Greenland estimate by 2.26 (308 000 km^2 / 136 000 km^2) to give the *K*-prior for the Barents Sea, i.e. approximated by a lognormal distribution with median of 1647 ktonnes and 95% confidence limits at approximately 679 and 5656 ktonnes: K-prior~lognorm(7.41,0.6). To allow for the added uncertainty in the approximate scaling of the West Greenland K-posterior to the Barents sea K-prior, we increased the CV by 10 percent points to 70% while maintaining the median at ca. 1647 tonnes which resulted in K-prior~lognorm(7.4,0.7).

4.7.2 Maximum Sustainable Yield (MSY)

A low information prior was given to this parameter in the form of wide normal distribution truncated at the lower end at 20 ktonnes: MSY~dnorm(70,0.000156) I(20,). this distribution is in practice uniform between 20 and 120 ktonnes and slowly declining thereafter. 20 ktonnes is the lowest annual catch registered and the stock has persisted over extended periods with substantially higher catches and therefore not likely that MSY is lower than 20 ktonnes. Catches above 120Kton is outside the catch history and MSY values higher than that is considered less likely.

4.7.3 Initial relative biomass (B_0/B_{msy})

There is little information about the pristine size of a stock i.e when the records started. The ″initial″ relative stock biomass in 1970 is considered to have been high as the fishery at that time was confined to inshore areas only. This parameter was given a normal distribution with mean = 1.5 and sigma = 0.26, i.e. a wide distribution with a mean between K and B_{MSY} under the assumption of a symmetrical Schaefer production curve.

4.7.4 Observation and process error

The error terms (CV's) for the input dataseries were given a gamma distribution with a 95% range of 10-30%, thought to be the typical range for such data and to give each dataseries equal prior weighting. The process error (CV) was uniform between .05 and .5 as we had no prior knowledge of the value of that parameter.

4.7.5 Catchability or scaling parameters (q)

The catchability coefficients, qR, qC, qRu and qE, interact with the carrying capacity, K. A uniform distribution was therefore not non-informative, and a prior distribution uniform on a log scale was preferred as a reference prior (cf. Gelman *et al.*, 1995, Punt and Hilborn, 1997; McAllister and Kirkwood, 1998) (Table 4.7.1). Their truncations were chosen wide enough not to interfere with the posterior and not to extend a stock estimate far beyond what a potentially maximum or minimum value for K.

4.7.6 Shape parameter (n)

In previous assessments this parameter was fixed at 2 i.e. assuming a symmetric "Schaefer production curve". Input data has little information on what this parameter should be, and we have little information otherwise whether the production curve for shrimp should be skewed in one or the other direction. Therefore, and to promote model stability this setting was kept in the final model.

Table 4.7.1. Priors used in the model implemented in OpenBugs (upper) and SPiCT (lower). ~ means "distributed as..", dunif = uniform-, dlnorm = lognormal-, dnorm= normal- and dgamma = gamma distributed. For the OpenBugs version the second parameter of the distributions are the precision while in the SPiCT implementation it is standard deviation.

OpenBugs

Parameter		Prior	
Name	Symbol	Type	Distribution
Maximal Suatainable Yield	*MSY*	low-informative	~dnorm(70,0.000156)
Carrying capacity	*K*	informative	~dlnorm(7.12,2.6)
Catchability survey 1	q_R	reference	ln(q_R)~dunif(-2,1)
Catchability survey 2	q_{Ru}	reference	ln(q_E)~dunif(-2,1)
Catchability survey 3	q_E	reference	ln(q_E)~dunif(-2,1)
Catchability CPUE	q_C	reference	ln(q_C)~dunif(-8,-5)
Initial biomass ratio	P_0	informative	~dnorm(1.5,15)
Precision survey 1	$1/s_R^2$	informative	~dgamma(4,0.1125)
Precision survey 2	$1/s_{Ru}^2$	informative	~dgamma(4,0.1125)
Precision survey 3	$1/s_E^2$	informative	~dgamma(4,0.1125)
Precision CPUE	$1/s_C^2$	informative	~dgamma(4,0.1125)
Precision model	s_P	reference	~dunif(.05,0.5)
Shape of production curve	*n*	informative	constant=2

SPiCT

Parameter		Prior	
Name	Symbol	Type	Distribution
Maximal Suatainable Yield	*MSY*	-	-
Carrying capacity	*K*	informative	lognorm(7.12,0.7)
Catchability survey 1	q_R	-	-
Catchability survey 2	q_{Ru}	-	-
Catchability survey 3	q_E	-	-
Catchability CPUE	q_C	-	-
Initial biomass ratio	$P_0/2$	informative	lognorm(0.75, 0.25)
CV survey 1	s_R^2	-	-
CV survey 2	s_{Ru}^2	-	-
CV survey 3	s_E^2	-	-
CV CPUE	s_C^2	-	-
CV model	s_P	-	-
Shape of production curve	*n*	informative	constant=2

4.8 Length information and recruitment index

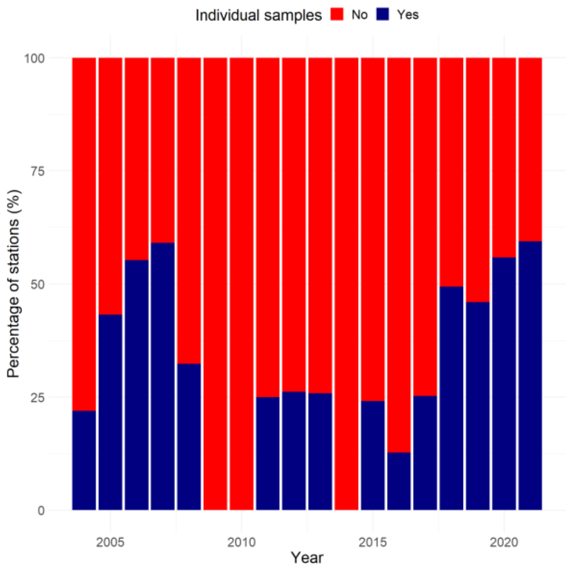
Estimating a recruitment index could improve stock estimates and forecasts by providing information on year-class strength, possibly allowing for a transition from a surplus production model to a stage- or length-based model. The development of a recruitment index requires sufficient data on length- or stage-composition of the stock. Earlier investigations (Hvingel, 2006) had found such data to be inadequate. However, available data on length and stage of shrimps from the Barents Sea ecosystem survey (BESS), Demersal fish survey in winter (WS) and commercial data were at this instance re-examined, to evaluate their suitability for estimating a recruitment index.

The analysis confirmed substantial gaps in data availability, both lack of spatial coverage as well as an overall low proportion of shrimp catches with individual shrimp sampling in some years (Figure **Error! Reference source not found.**). The BESS represents the best data source, both in terms of quantity and quality. Nevertheless, the sample numbers and representativeness has been insufficient, especially in the period between 2009 and 2017 when individual samples were collected at less than 25% of all stations, including no individual samples at all in 2009, 2010, and

2014. In addition, the available data is heavily skewed towards the western part of the Barents Sea. The reasons for this lack of sampling coverage are inconsistencies shrimp sampling protocols in earlier years and technical issues regarding data exchange between Norway and Russia. Data on length composition from the WS and commercial catches is patchy at best, as sampling efforts have been limited.

The WK further notes that the potential for a full age/cohort-based segregation of survey data will in any case be difficult due to the relative slow growth of the Barents Sea shrimp, e.g. compared to the North Sea stock. The resulting absence of clear modes in the length distribution prevent the use modal analyses to separate length modes and cohorts (Figure).

Individual sampling has improved in recent years, and in 2021 a new sampling protocol was implemented that aims at ensuring representative length and stage information across the survey area with the available resource. In addition, the need to include individual shrimp data in Norwegian-Russian data exchange was reemphasized and has resulted in clear improvements in 2021. It can therefore be expected that the quality of length and stage distribution data will be acceptable in the future. The conclusions at WKPAND22 were therefore that i) *current dataseries on length and stage distributions are insufficient for the estimation of a representative recruitment index*; ii) *for the future it is recommended at regular intervals to revisit this issue to ensure that data quality is sufficient to potentially found the estimation of recruitment indices.*



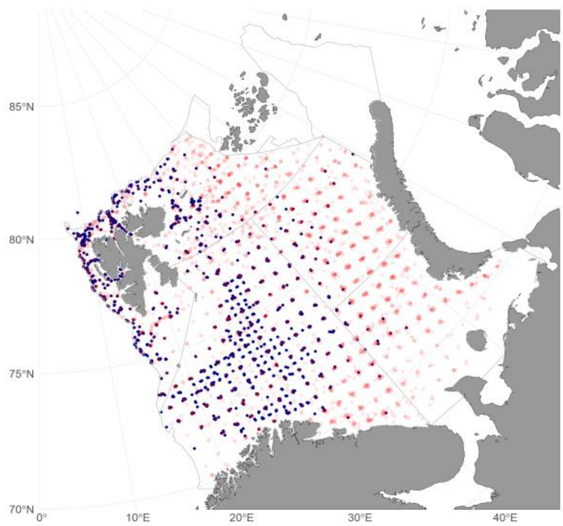


Figure 4.8.1. Proportion of stations with individual shrimp samples compared to all stations with shrimp catches at the BESS (left) and distribution of station with and without individual samples since 2004 (right).

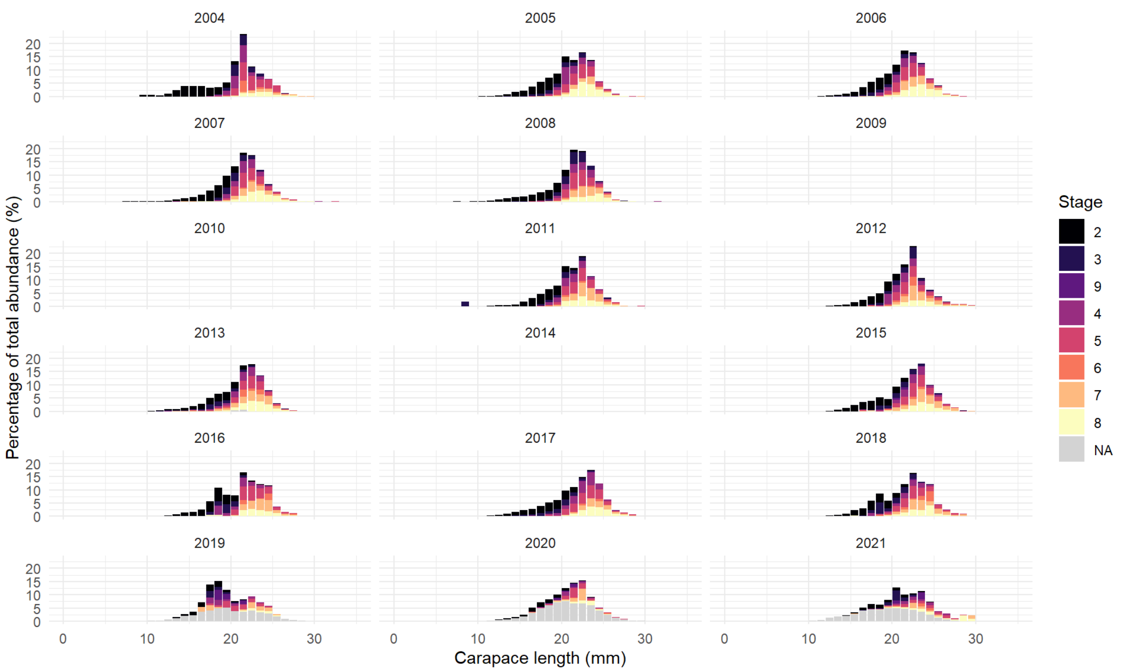


Figure 4.8.2. Length-stage frequency distribution of shrimp caught during the BESS. Shrimp that were measured but not staged are shown in grey.

4.9 Assessment model

Philosophically the two estimation frameworks (OpenBugs and SPiCT) are different and long discussions may be found in the literature addressing Bayesian vs frequentist approach. Thinking about the true estimate of a model parameter e.g. MSY as a probability distribution of a certain concave shape, SPiCT will focus on estimating the mode of this distribution and then construct by some approximation the remaining density distribution. OpenBugs estimates the entire distribution directly.

The OpenBugs-based model has been used for the stock assessment until 2021 and represents therefore the baseline. While the OpenBugs implementation is more flexible in most respects (e.g. can implement time-series of estimated CVs of the observation series), SPiCT on the other hand provides advantages through being a published surplus production model framework that is widely tested and used within ICES to assess a range of stocks. In a first step, implementations of both modelling frameworks were created in R and compared against each other to ensure sufficiently identical results with comparable configurations. Based on the results, SPiCT was selected as modelling framework and after an initial exploration, candidate models were defined and tested. Lastly, sensitivity analysis was conducted around the selected reference model to test for the robustness of the model in respect to specific configurations, notably the priors and the input time series.

4.9.1 Comparison of assessment frameworks

When set up approximately equally (as technically possible) with regards to input data and priors, the two implementations give similar results (Figure 4.9.1). It was therefore concluded unlikely that selection of any of the two methods of implementation would significantly influence assessment results.

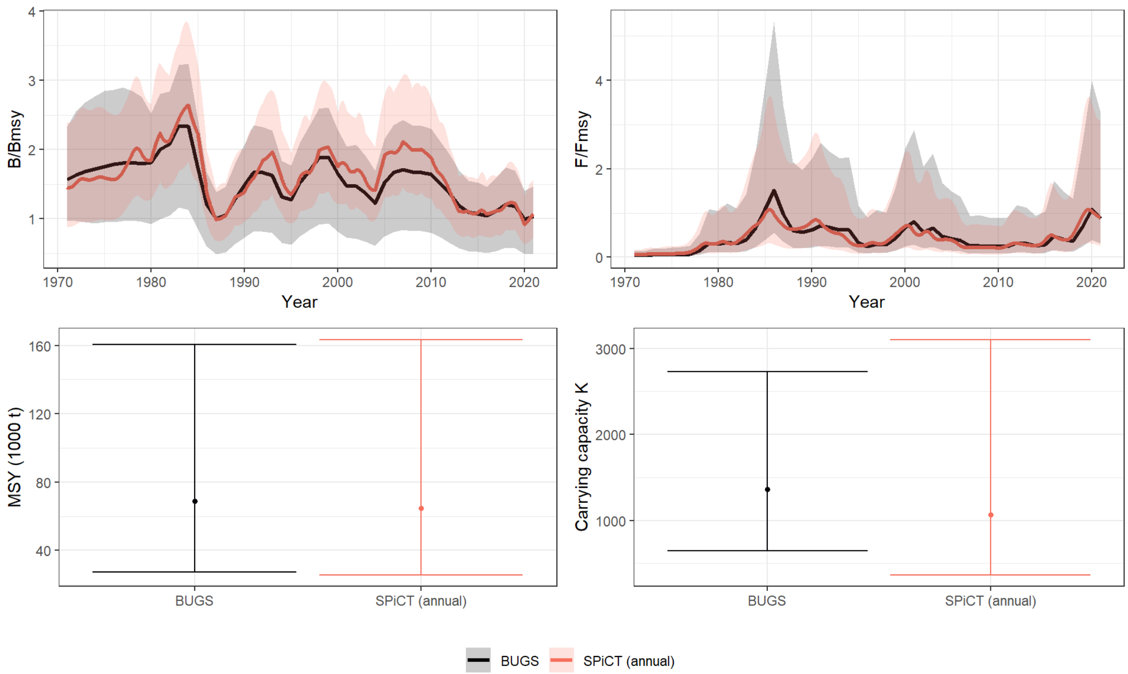


Figure 4.9.1. Assessment summaries (B/Bmsy, F/Fmsy, MSY, K) of BUGS and SPiCT runs with proposed configurations and input data.

4.9.2 Effects of revised time-series and priors

The transition from the 2021 assessment to the WKPRAWN 2022 revised framework, results in a change in the resent stock trajectory, but also in model parameters the define stock dynamics (Figure 4.9.2). This is mainly due to the updated stock-index-series, which overall are more pessimistic for the recent time-period. The revised input priors (the more informative K-prior in particular) influence assessment results in the same direction. In conjunction, these two revisions produce a current stock status closer to, but likely still above, B_{msy} than seen in the 2021 assessment and a production potential, MSY, at a lower and more plausible magnitude (Figure 4.9.2).

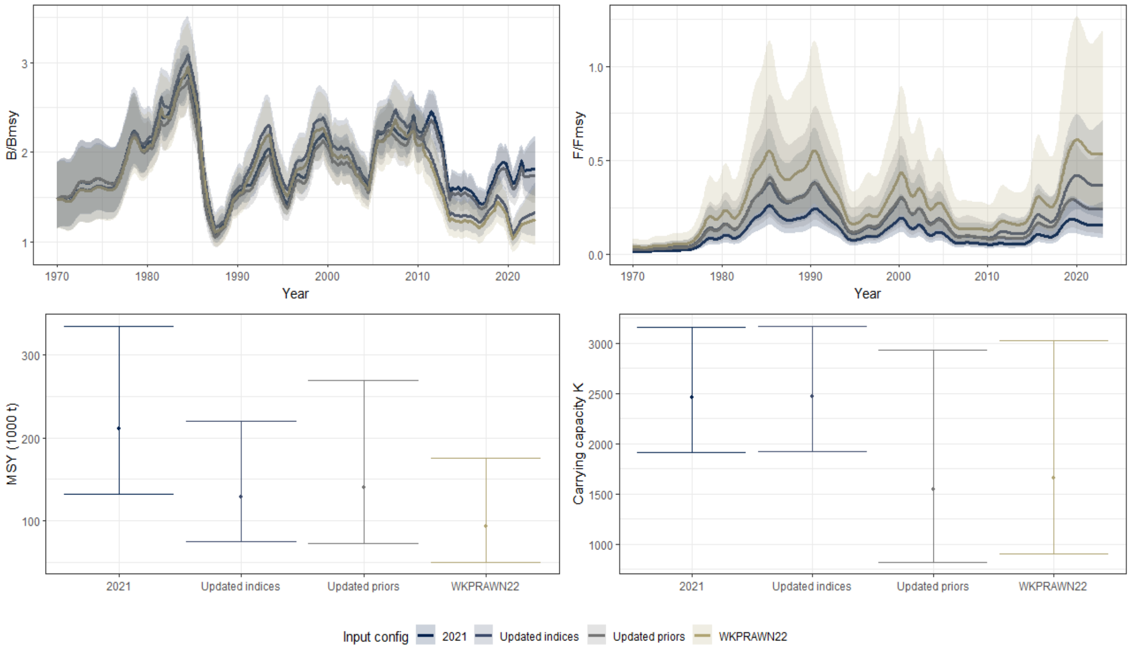


Figure 4.9.2. Estimated B/B$_{msy}$ and F/F$_{msy}$ time series and corresponding MSY and F$_{msy}$ for SPiCT runs using the configuration as reported in the 2021 assessment, compared to versions with updated indices and/or priors as proposed at WKPRAWN22.

4.9.3 SPiCT stock assessment

The stochastic surplus production model in continuous time (SPiCT) implements surplus production models as continuous-time state-space model, with biomass and fishing mortality modelled as unobserved process and both catch and stock indices as observations with observation error (Pedersen & Berg, 2017). Within ICES, SPiCT has become a recommended solution for stocks where data are insufficient for age- or length-structured models, initially for category 3 stocks (Berg *et al.*, 2021) and subsequently also an option for category 2 or 1 stocks (Mildenberger *et al.*, 2021). Here we used to the most recent (1.3.5) R-implementation of SPiCT.

4.9.4 Candidate models

Four candidate models were defined based on the *a priori* configuration of the model and exploration: 1. the reference model with the predefined priors for K, B0/K, the shape fixed to a Schaefer model, default Euler approximation (dteuler = 1/16), and four input time series (3 survey indices, 1 CPUE index; 2. the reference model without K prior; 3. the reference model without the CPUE index included; and 4. a discrete form (dteuler = 1) of the reference model).

The candidate models were evaluated based on the criteria for the acceptance of a SPiCT model (Mildenberger *et al.*, 2021). The results show that no mode fully satisfied all criteria, but both the reference model and the discrete version showed only minor violations. Specifically, there was significant autocorrelation in OA residuals of survey indices in all cases. Furthermore, there was slightly larger uncertainty than one order of magnitude for F/F$_{MSY}$ and a minor proportion of convergence issues and alternative estimates when using jittering on the initial parameter values. Removing the K prior or CPUE index caused a deterioration of model performance, increasing the uncertainty of B/B$_{MSY}$ and F/F$_{MSY}$ substantially and resulting in strong retrospective patterns. The model without K prior performed the worst, becoming much less stable than when using the K prior.

It was therefore decided to accept the reference model as the new assessment model configuration. The following sections detail outputs and diagnostics of the reference model. Detailed outputs of the other candidate models and further information can be found in the annex.

Criteria	Reference	no K prior	no CPUE index	discrete
Convergence	Yes	Yes	Yes	Yes
Variance pars finite	Yes	Yes	Yes	Yes
No OAR violations	Minor	Minor	Minor	Minor
Consistent retros	Yes	No	No	Yes
Realistic production curve	Yes	Yes	Yes	Yes
Uncertainty < 1	Partially	No	No	Partially
No effect of initial values	Partially	Partially	No	Partially

4.9.5 Reference model

The configuration used in the reference model is detailed in Table 4.9.1. A summary of the model output is shown in Figure .

Table 4.9.1. Configuration of reference SPiCT model.

Data/parameter	Settings
Catch data	Total landings 1970–2021, by year
Standardized indices	Norwegian shrimp survey 1982–2004 (timing +0.5) Russian survey 1984–2002 (+0.5) Barents Sea ecosystem survey 2004–2021 (+0.75) CPUE index 1980–2021 (+0.5)
Prior: logn	Log(2),0.001
Prior: logK	Log(1647), 0.7
Prior: logbkfrac	Log(0.75), 0.25
Time step: dteuler	1/16

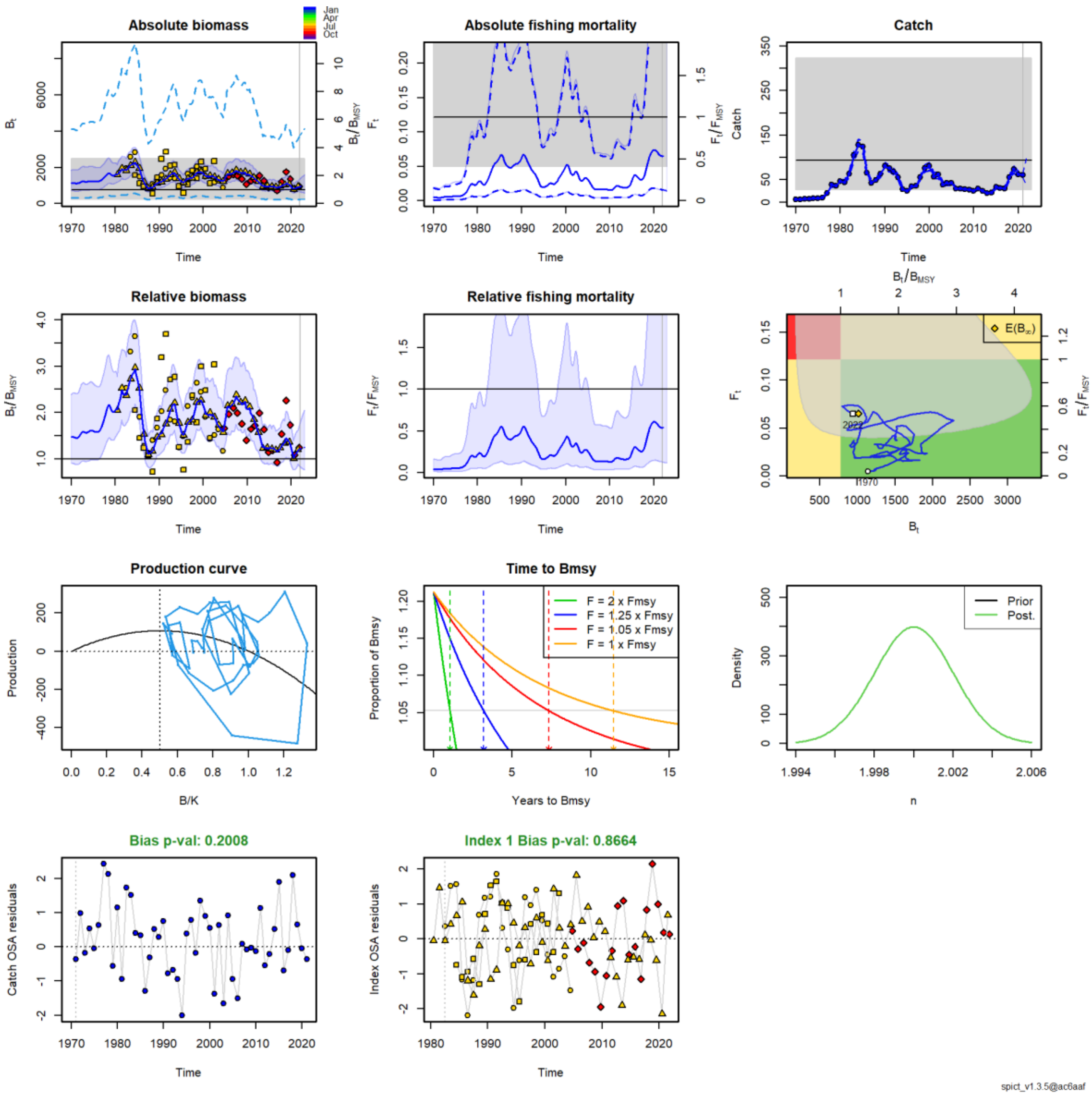


Figure 4.9.3. SPiCT output of the reference model using updated indices and priors as proposed at WKPAND22.

4.9.6 Model diagnostics and retrospective patterns

Model diagnostics showed no major violation in the reference model. One-step ahead residuals were well distributed and did not show any bias (Figure). There were minor violations in terms of significant autocorrelation in the residuals of the two old survey indices, due to the model ignoring them largely. Retrospective analysis resulted in very stable patterns for B/B_{MSY} while there were slightly stronger deviations in F/F_{MSY} (Figure 4.9.5). However, Mohn's rho for F/F_{MSY} was below the critical threshold of -0.22 for short-lived species (Hurtado-Ferro, 2014). Based on the presented diagnostics, the reference model was considered as adequate and acceptable.

The model output showed a strong correlation between K and the index catchabilities q (-0.98). This may seem counterintuitive considered the high precision of B/B_{MSY}, since in model formulation used here B/B_{MSY} is a direct function of K and q. However, as K and q are directly linked, the join term can essentially absorb the negative correlation (high q and low K can give similar value to low q and high K value), resulting in narrow confidence limits on the estimated B/B_{MSY}. This may result in precise but biased estimates of B/B_{MSY}, as observed in some stocks (Bouch *et al.*, 2020). The level of contrast in the data in the present assessment may thus potentially bias and challenges in accurately estimating B/BMSY and F/FMSY. However, the informative prior on K helps to anchor the model and is therefore critical for model stability (as confirmed by the candidate model without K prior). The work to define a suitable but not too narrow (see also 4.9.7 Sensitivity analysis) prior for K has therefore been important for the acceptance of the SPiCT model. In contrast to B/BMSY, the imprecision of F/FMSY could not be fully resolved and should therefore be further investigated in the future.

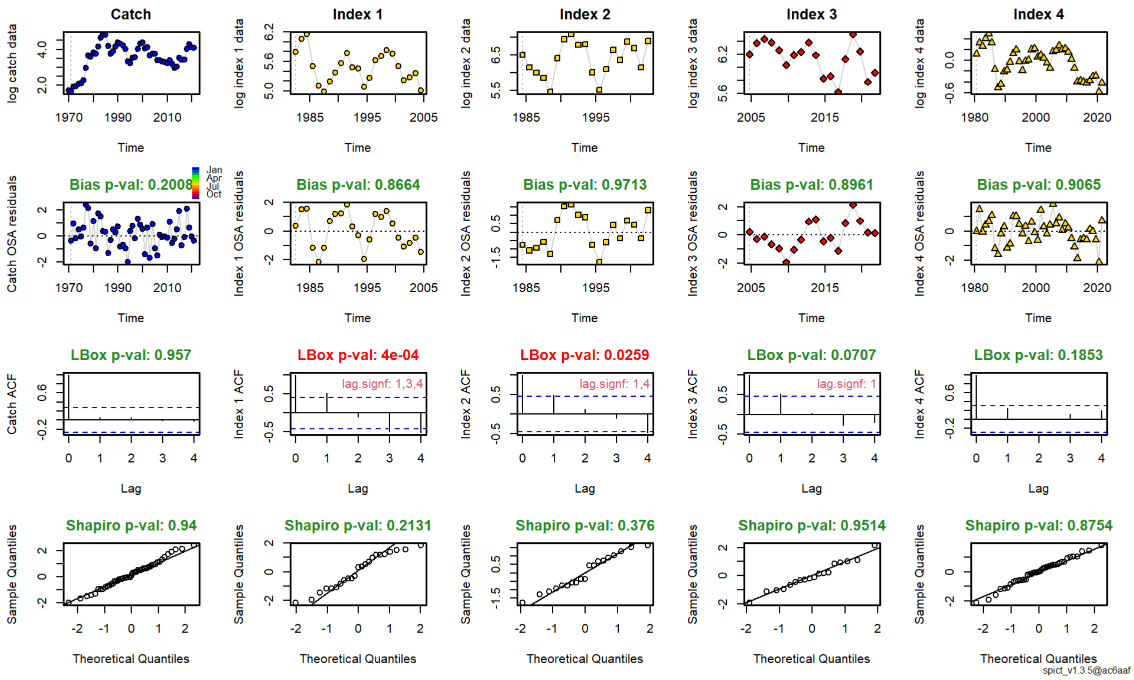


Figure 4.9.4. SPiCT model diagnostics of the reference model as proposed at WKPAND22.

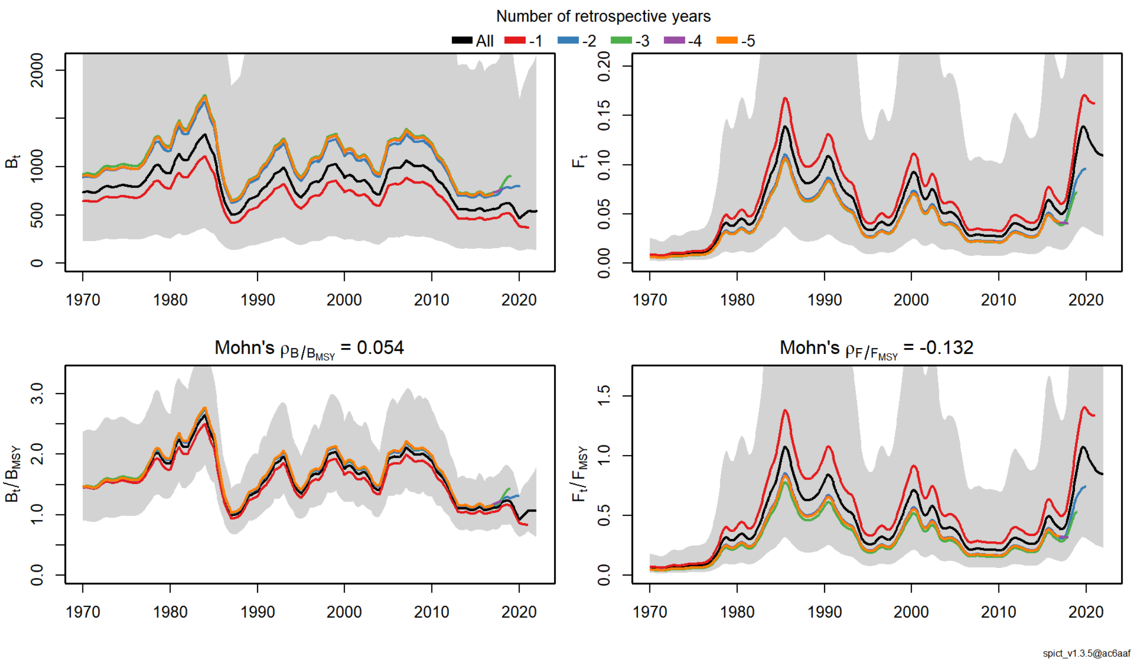


Figure 4.9.5. Retro runs of SPiCT reference model with the configuration proposed at WKPAND22.

4.9.7 Sensitivity analysis

Sensitivity analysis was conducted to explore the effects of selected input data and prior on the estimated trends and parameters. In general, some of the strongest effects were observed when excluding the CPUE index (Figure 7), whereas the exclusion of any of the three survey indices had little to no effect. This highlights that given the current input data, the CPUE index drives the stock assessment model to a large degree. This can be explained by both the length of the CPUE time series that spans almost the entire assessment period compared to the more fragmented survey indices, and by the relative smoothness of the CPUE compared to the larger variation in the survey indices. Subsequently, the model estimates tend to follow closely the CPUE, ignoring some of the additional fluctuations indicated by the survey indices.

Sensitivity to prior definitions was limited and followed the expectations. Estimated K scaled with the mean of K prior, however less than proportionally (Figure). Other parameters were less sensitive to even large deviations (±50%) of prior mean, just resulting in a minor scaling effect. This underlined that while the K prior stabilized the model and was thus considered as essential, it did not have any outsized impact on the estimated trends and reference points. The initial depletion prior (B_0/K) was only relevant for the initial period where no index information is available but had negligible effects forward in time (Figure). Similarly, the assessment model was insensitive to the definition of the production model shape, here fixed to a Schaefer model, as more relaxed priors or default settings had only minor impacts Figure).

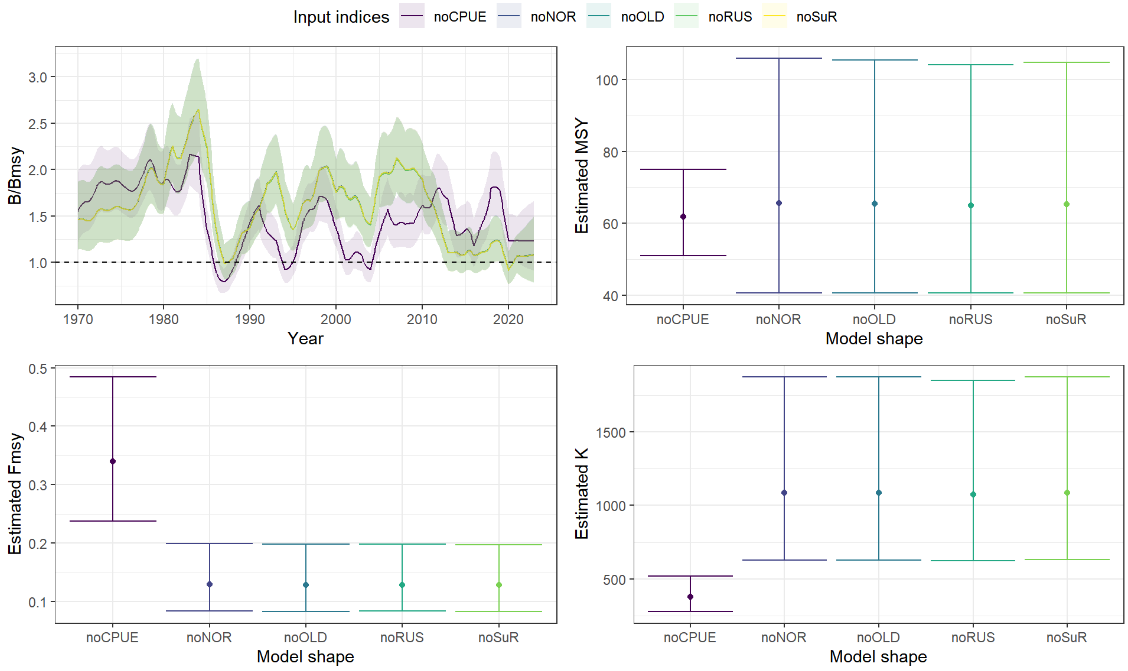


Figure 7. Sensitivity of SPiCT estimates to removing standardized indices from full model, using otherwise the model configuration proposed at WKPRAWN22.

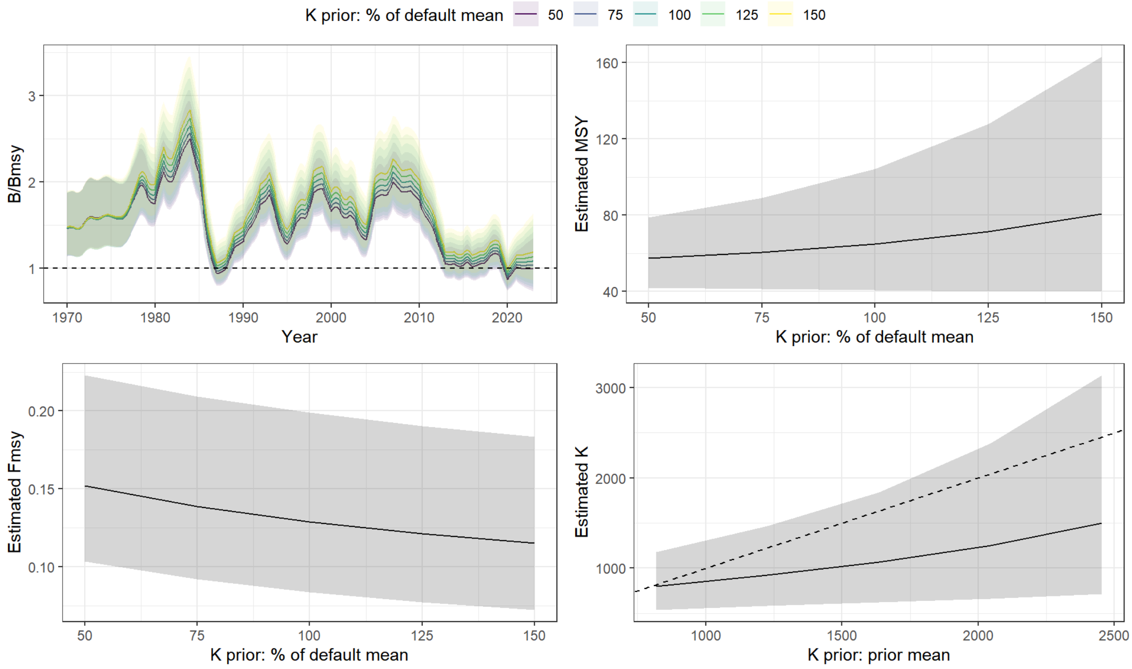


Figure 4.9.7. Sensitivity of SPiCT estimates to carrying capacity prior mean, using the model configuration proposed at WKPRAWN22.

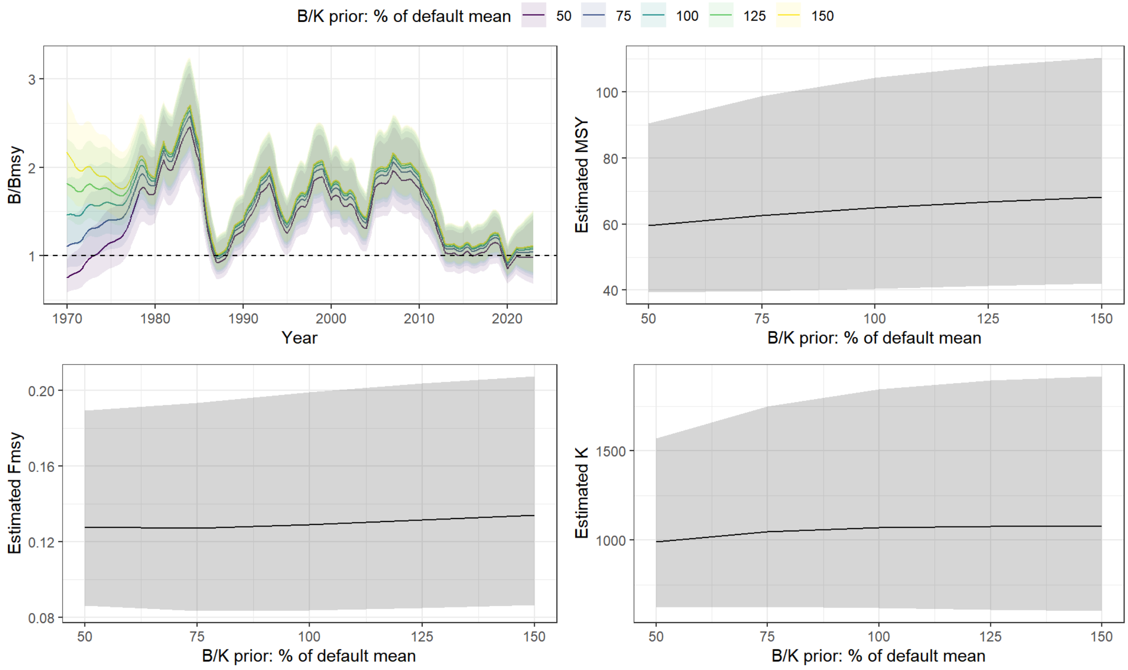


Figure 4.9.8. Sensitivity of SPiCT estimates to initial depletion (B0/K) prior mean, using the model configuration proposed at WKPRAWN22.

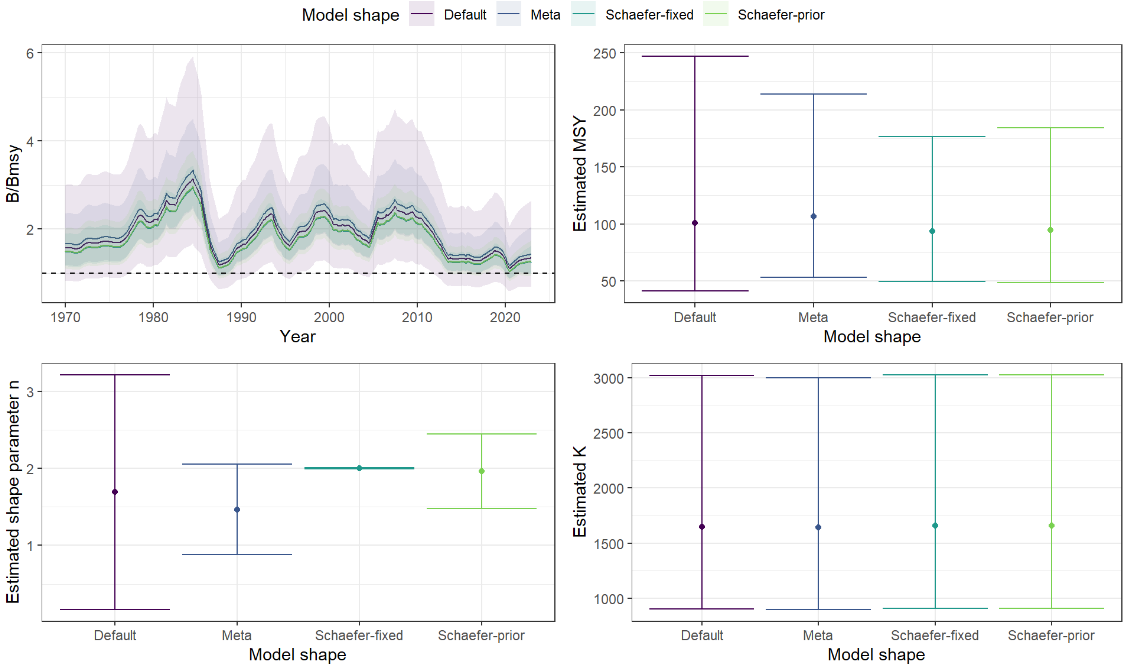


Figure 4.9.9. Sensitivity of SPiCT estimates to production model shape, compared the reference model configuration proposed at WKPRAWN22 with a fixed Schaefer model shape with a Schaefer prior, a prior from a meta study (Thorson *et al.*, 2012), and default SPiCT prior.

4.10 Short-term projections

The assessment is done using data for the prior years and the partial data for the assessment year. The short-term forecast to the end of the assessment year is based on status quo fishing mortality, potentially modified according to available information from the industry.

In the absence of an explicitly defined Harvest Control Rule (HCR) by managers, the default HCR applied, is "a hockey-stick" with break points at *Btrigger* and *Blim* (ICES 2019). The TAC advice is the 35th percentile of the short-term forecast of the catch distribution corresponding to F at F_{MSY}. When the biomass is less than $B_{trigger}$, F is reduced linearly to zero at B_{lim}. In SPICT this is equal to the option "8. ICES advice rule".

Alternative management options are derived as per standard SPiCT output and may be included in the management advice, as the expert group and ICES finds relevant:

1. **currentCatch**: Keep the catch of the current year (i.e. the last observed catch)
2. **currentF**: Keep the F of the current year
3. **Fmsy**: Fish at Fmsy i.e. $F=F_{msy}$
4. **noF**: No fishing, reduce to 1% of current F
5. **reduceF25**: Reduce F by 25%
6. **increaseF25**: Increase F by 25%
7. **msyHockeyStick**: Use ICES MSY hockey-stick advice rule
8. **ices**: Use ICES MSY 35th hockey-stick advice rule

4.11 Appropriate Reference Points

There were no changes made to the reference points introduced in the assessment of this stock in 2006 (B_{lim} and F_{lim}) and 2010 ($B_{trigger}$). Stock biomass (B) and fishing mortality (F) is measured on a relative scale such that:

- Relative_F is F_y/F_{MSY}, where F_y is the estimated F in a year and F_{MSY} is the F that maximizes yield.
- Relative_B is B_y/B_{MSY}, where B_y is the estimated biomass in a year and B_{MSY} is the biomass corresponding to MSY

In accordance, the reference points are:

- Relative_$B_{trigger}$= 0.5, Relative_B_{lim}=0.3 and Relative_F_{lim}=1.7.

4.12 Future Research and data requirements

- Management units: Investigations into the existence of biological stock structure should continue e.g., through further evaluation of genetic information. Also, and in connection with a potential partition of the current single stock management unit in two (e.g. inshore/offshore units), the implications of different fleets targeting different parts of the population should be analysed. Ongoing work aims at establishing a dedicated stock assessment for coastal shrimp Northern Norway, which would add further justification for separate management units.
- Cod predation: For other shrimp stocks, cod (*Gadus morhua*) has been shown to be an important predator with potential to influence shrimp stock dynamics. Current analyses could be refined to include spatial and size (both shrimp and cod) and to include available data from cod stomach sampling.
- Recruitment index: An improved sampling and data sharing protocol has been established, likely resulting in better individual data in the future. A recruitment index should therefore be revisited in the future.
- Fishery data: The shrimp fishery in international waters in the Barents Sea represents about 1/3 of the total catches. The assessment could benefit for better catch and effort data from this fishery.
- Index uncertainty in the assessment model: Inclusion of estimated index uncertainty was tested during the exploration of the SPiCT model, with a relevant impact on estimates (
- Figure). There is large variation in the inter-annual uncertainty of the indices (linked to e.g. incomplete coverage), making the explicit incorporation of index uncertainty a promising avenue that should be further investigated.
- Seasonal dynamics in the assessment model: the continuous time formulation of SPiCT allow for the modelling of explicit seasonal dynamics. The shrimp fishery in the Barents Sea has a clear seasonal pattern and the inclusion of this information should therefore be explored.
- Using a WS index or combined BESS-WS index in the assessment: as part of WKPRAWN22, the estimation of a separate WS index or combined BESS-WS index was tested to utilize the existing WS data despite its prevalent coverage issues. Although it was decided to not include this information at the current point due to some remaining issue, it was generally considered as a worthwhile approach that could help to integrate further seasonal information. A comparison is shown in Figure 4.12.2.

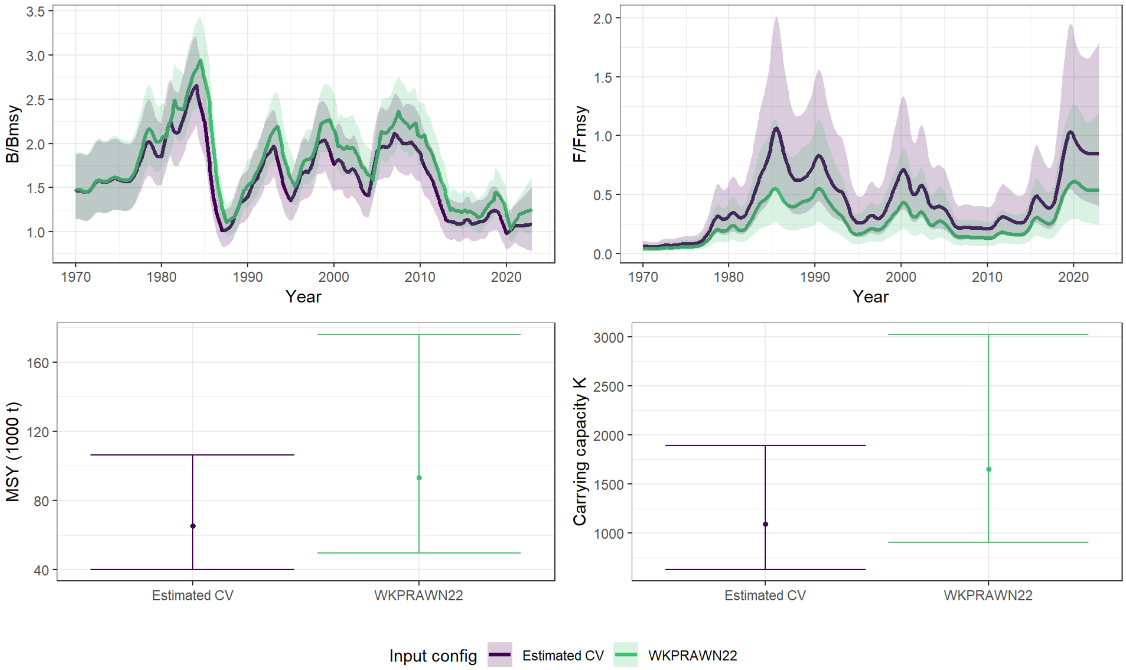


Figure 4.12.1. Estimated trends and parameters from the reference model (WKPAND22) and an altered version where annual uncertainty of standardized indices was weighted by the estimated index uncertainty.

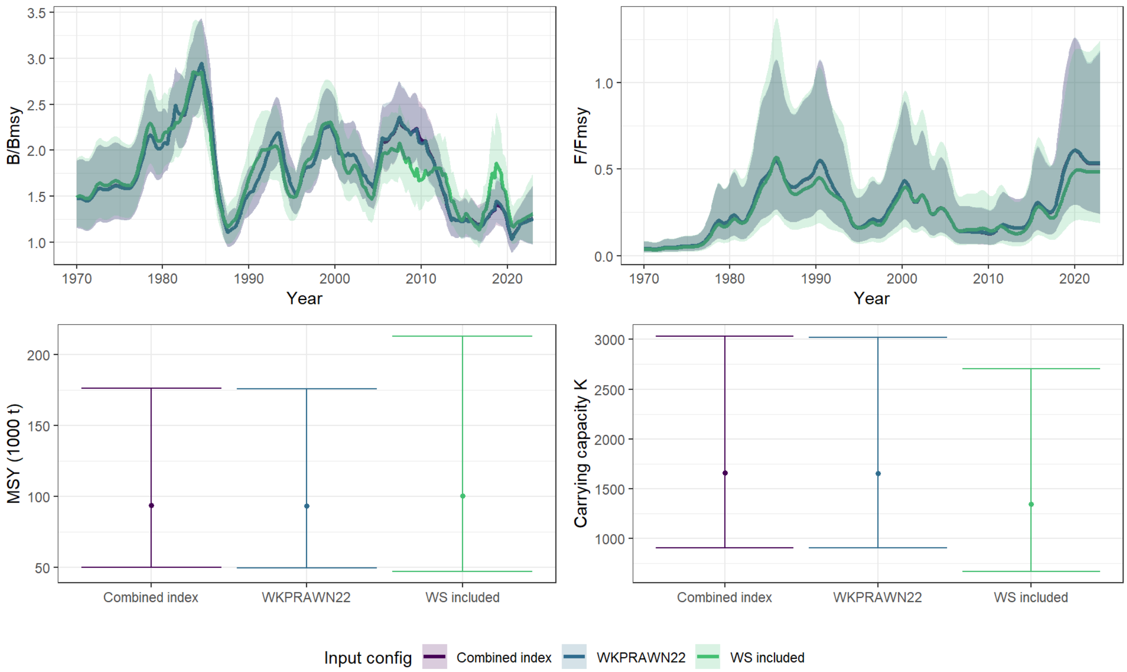


Figure 4.12.2 Estimated trends and parameters from the reference model (WKPAND22) compared to alternative versions that include the a combined BESS-WS index instead of the selected BESS index or a WS index in addition to the BESS index.

4.13 References

Anderson, S.C. 2019. A reproducible data synopsis for over 100 species of British Columbia groundfish. Canadian Science Advisory Secretariat (CSAS).

Aschan, M., Johannesen, E., Bogstad, B. and Hvingel, C. 2006. Why Does Not the Shrimp Stock in the Barents Sea Respond on Predation by Cod? NAFO SCR Doc. 06/71, Serial No. N5321.

Beguin, J., Martino, S., Rue, H., and Cumming, S.G. 2012. Hierarchical analysis of spatially autocorrelated ecological data using integrated nested Laplace approximation. Methods in Ecology and Evolution 3(5): 921-929.

Berg, C., Coleman, P., Cooper, A., Hansen, H.Ø., Haslob, H., Herrariz, I.G., Kokkalis, A., Mildenberger, T., Moura, T., and Pennino, M.G. 2021. Benchmark Workshop on the development of MSY advice for category 3 stocks using Surplus Production Model in Continuous Time; SPiCT (WKMSYSPiCT).

Bouch, P., Minto, C., and Reid, D.G. 2020. Comparative performance of data-poor CMSY and data-moderate SPiCT stock assessment methods when applied to data-rich, real-world stocks. ICES Journal of Marine Science 78(1): 264-276. doi:10.1093/icesjms/fsaa220.

Breivik, O.N., Aanes, F., Søvik, G., Aglen, A., Mehl, S., and Johnsen, E. 2021. Predicting abundance indices in areas without coverage with a latent spatio-temporal Gaussian model. ICES Journal of Marine Science 78(6): 2031-2042. doi:10.1093/icesjms/fsab073.

Brooks, M., Kristensen, K., van Benthem, K., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Maechler, M., and Bolker, B. 2017. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal 9(2): 378-400. Available from https://journal.r-project.org/archive/2017/RJ-2017-066/index.html. [accessed.

Burmeister, A. and Rigét, F.F. 2020. The West Greenland trawl survey for Pandalus borealis, 2020, with reference to earlier results. NAFO SCR Doc. 20/053, Serial No. N7127.

Gelman, A., Carlin, J.C., Stern, H. and Rubin D.B. 1995. Bayesian Data Analysis. Chapman and Hall, New York. 525 pp.

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., and Muradian, M.L. 2014. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science: Journal du Conseil 72((1)): 99–110. doi:10.1093/icesjms/fsu198.

Hvingel, C. 2006 A. Construction of biological advice for the management of a northern shrimp fishery – the West Greenland example. Thesis Dr. scient. Norwegian College of Fishery Science, University of Tromsø, Norway. 72 pp.

Hvingel, C. 2006 B. Towards a Quantitative Assessment Framework for the Shrimp (Pandalus borealis) Stock in the Barents Sea. NAFO SCR Doc. 06/64, Serial No. N5314.

Hvingel, C., Kingsley, MsC. 2006. A framework to model shrimp (Pandalus borealis) stock dynamics and to quantify the risk associated with alternative management options, using Bayesian methods. ICES Journal of Marine Science 63 p. 68-82. Doi: https://doi.org/10.1016/j.icesjms.2005.09.002.

McAllister, M.K. and Kirkwood, G.P. 1998. Bayesian Stock assessment: a review and example application using the logistic model. ICES Journal of Marine Science, 55: 1031–1060.

Mildenberger, T., Kokkalis, A., and Berg, C. 2021. Guidelines for the stochastic production model in continuous time (SPiCT).

NAFO/ICES. 2020. Report of the NAFO/ICES Pandalus Assessment Group Meeting, 26 - 30 October 2020, WebEx. NAFO SCS Doc. 20/21.

Pedersen, M.W., and Berg, C.W. 2017. A stochastic surplus production model in continuous time. Fish and Fisheries 18(2): 226-243. doi:https://doi.org/10.1111/faf.12174.

Punt, A. and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries, 7: 35–63.

Shumway, S.E., H.C. Perkins, D.F. Schick, and A.P. Stickney. 1985. Synopsis of biological data on the pink shrimp, Pandalus borealis Krøyer, 1838. NOAA Technical Report NMFS 30. FAO Fisheries Synopsis No. 144. 57 pp.

Thorson, J. T., Cope, J. M., Branch, T. A., Jensen, O. P., and Walters, C.J. 2012. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. Canadian Journal of Fisheries and Aquatic Sciences 69(9): 1556-1568. doi:10.1139/f2012-077.

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

5 pra.27.3M Flemish Cap

5.1 Stock ID and sub-stock structure

No new information on stock structure was discussed.

5.2 Issue list

The Issue list for pra.27.3M is given below.

Stock name	Northern shrimp (*Pandalus borealis*) on the Flemish Cap (NAFO Div. 3M)	
Stock code	Northern shrimp in Div. 3M	
Stock coordinator	Name: Jose Miguel Casas	Email: mikel.casas@ieo.es
Stock assessor	Name: Jose Miguel Casas	Email: mikel.casas@ieo.es
Data contact	Name: Jose Miguel Casas and Kalvi Hubel	Email: mikel.casas@ieo.es, kalvi.hubel@ut.ee,

Issue	Problem/Aim	Work needed / possible direction of solution	Data needed to be able to do this: are these available / where should these come from?	Responsible expert from WG	External expertise needed at benchmark type of expertise / proposed names
(New) data to be Considered and/or quantified					
Catch/Landings data	Historical catches by month, quarter or annual (1993-2021)	Compilation of input data from NAFO Secretariat (STATLAND) and from national data contacts (yearly).	Data already available al national level. NAFO Database (STATLANT 21) Also, NIPAG estimations.	Mikel Kalvi?	No
Discards	Estime discards back in time (1993-2020)	Compilation of historical data by specific studies.	Bibliographic Review	Mikel Kalvi?	Country representatives
Lengths data	Collect historical length data from commercial fishery	Compilation of input data by national data contacts, mainly Iceland and Estonia	Input data until 2006 comes mainly from Iceland sampling program. Others sources to be consulted should be Canada,	Mikel Kalvi?	No

Issue	Problem/Aim	Work needed / possible direction of solution	Data needed to be able to do this: are these available / where should these come from?	Re-sponsi-ble ex-pert from WG	External exper-tise needed at benchmark type of exper-tise / proposed names
			Estonia, Spain, Greenland and Russia.		
Age data	Collect historical age data from commercial fishery and research surveys	Review the histori-cal series.	Data partialy available from NIPAG SCR and WD documents	Mikel	
Recruitment Index	Collate available data on length and age composi-tion from survey and commercial catches	Review the histori-cal series	Length and age data available for UE surveys, but only age data for commercial fishery	Mikel	
Tuning se-ries	Revision of the historical CPUE international data series (1993-2009) + in-clusion of the new data for the resumed fishery and the technical creep after 9 years of morato-rium.	Historical review of the historical catches and effort to correct the wrong allocation of the cacthes be-tween 3M and 3L Divisions. Review the method of standardization of CPUE with the in-clusion of new data.	Data already available from Canada, Iceland, Faroes, Greenland, EU countries (Estonia, Spain), Russia, Norway	Mikel Kalvi?	
	Collate the EU Flemish Cap Survey series (1988-2021) and Faroese Flem-ish Cap Survey (1997-2003)	Review the histori-cal series indexes from Faeroes and UE	Data already available	Mikel	
Biological Parameters	Priors for stock assess-ment models	Explore the neces-sary priors for the proposed assess-ment models	M assumption could be theoretical. Sex ratio by length, length/weight relation-ship, is available from EU surveys.	Mikel	
Assessment method	Explore possible assess-ment method based in the available information e.g. length/stage based models (SS3), SPICT.., to determine required input data and possibility of building up necessary time series	Assess existing al-ternative assess-ment models, their suitability and data needs.	Catches, CPUE from commercial fishery and abundances and bio-mass from surveys se-ries are available. No data from 2011 to 2019 because de mora-torium.		NIPAG repre-sentatives ??
Biological Reference Points	Only B_{lim} has been de-fined	Progress in the quantitative assess-ment to define a fishing mortality Reference point		??	Yes

5.3 Ecosystem drivers

No new information was available. It was noted that there has been some research on ecosystem model for the Flemish Cap in recent years that could be very relevant for this stock.

5.4 Fishery dependent data

Catch data

Compilation of historical input catch data (1993-2021) are yearly carried out from NAFO Secretariat and from national data contacts. There were no catches from 2011-2019 because the moratorium. In 2022 a new moratorium was established.

In the years before the moratorium, mainly from 2007 there was serious concern about the misallocation of the catches between NAFO Divisions 3M and 3L. The last compilation catch data were presented during NIPAG meeting in SCR Doc 21/038.

Effort and CPUE data

Historical CPUE international data series (1993-2009) is available from Canada, Iceland, Faroes, Greenland, EU countries (Estonia, Spain), Russia, Norway. To solve the wrong allocation between 3M and 3L catches, from 2007 only the trips carried out exclusively in NAFO Div. 3M were considered. In 2020 and 2021 there was no available individual effort and catch data. Data are available in SCR 10/064

Also, there was no progress about the effect of technical creep in the standardization of CPUE with the inclusion of new data.

There is very poor information about discards in shrimp fishery.

5.5 Survey data

The best information on shrimp in 3M comes from the EU summer research survey series from 1988 to the present. The survey provided fisheries independent biomass and abundance indexes, length distributions and age composition (by RMIX). Also, a recruitment index is estimated considering the abundance of age 2 as indicator of recruitment. The results of the survey are presented annually in NIPAG meeting as SCR Doc document.

Weights by length and by age, length at maturity or length at sex change are derived from survey. Although the age composition for the population is annually carried out by the length modal analysis, the uncertainty is very high in the ages older than 4 years.

5.6 Assessment model

No analytical assessment is available. Evaluation of stock status is based upon interpretation of commercial fishery information and research survey data. During this benchmark workshop one attempt was made to apply a statistical age-structured population modelling framework (Stock synthesis SS3), with the available information from 3M shrimp stock. However, the limited experience of the stock responsible with this kind of models and the lack of time to spend in these tasks, made it very difficult to progress adequately. Also, the performance of the evaluation presented serious difficulties due to the lack of relevant information (among others, length distribution from commercial vessel and priors for M that should be assumed theoretically).

5.7 Future Research and data requirements

SS3 is a very versatile software, but it can be complex and quite difficult to use effectively with a stock that has never been quantitatively assessed before.

One option for the future research it is to explore a simpler model such as the Stochastic Surplus Production model in continuous time (SPiCT), which incorporates dynamics in both biomass and fisheries.

6 Reviewer reports

Benchmark workshop on Pandalus stocks (WKPRAWN 2022)

Reviewers' Report: Cóilín Minto

Colm Lordan (Chair), Ewen Bell and Cóilín Minto served as the external reviewers for the WKPRAWN benchmark for the following northern shrimp (*Pandalus borealis*) stocks:

pra.27.3a4a Divisions 3.a and 4.a East (Skagerrak and Kattegat and northern North Sea in the Norwegian Deep)

pra.27.1-2 in subareas 1 and 2 (Northeast Arctic)

Flemish Cap (NAFO Div. 3M)

Summary

The Benchmark Workshop met virtually from 24-28 January 2022 and was attended by stock experts, assessment scientists, observers and the external reviewers. Stocks were first introduced in round table presentations to familiarise all with the stock details, followed by proposed assessments for each stock. During the course of the meeting additional data explorations and assessment runs were requested by both participants and reviewers, and the responsible scientists responded efficiently to these requests.

Major changes were proposed in the assessment methodologies for pra.27.3a4a and pra.27.1-2; namely: an ensemble of Stock Synthesis (SS3) models for pra.27.3a4a and a move to a Surplus Production in Continuous Time (SPiCT) model for pra.27.1-2. Major changes were also made to derive reference points for pra.27.3a4a reflecting a change in reference point basis; and to reference points for pra.27.1-2 reflecting a change in model and recent recommendations regarding SPiCT reference points. These are reviewed in detail by stock below.

No proposed assessment of the Flemish Cap stock was presented. An overview of the data was provided and is used as the basis for recommendations for future work on this stock presented below.

Time was relatively limited to assist inter-continental participation but resulted in a relatively fast pace for a lot of changes made. Draft working documents were uploaded on the Friday preceeding the workshop and did not provide sufficient time to review in advance, given the magnitude of the changes made. In future, it is recommended that appropriate lead-in time is provided to reviewers when reviewing such significant changes to the methods. Notwithstanding, the Benchmark meeting was conducted to a high scientific standard, the debate was thorough and constructive.

A review of the issues raised at the Benchmark Workshop is presented by stock in the sections that follow.

Northern Shrimp in Divisions 3.a and 4.a East (Skagerrak and Kattegat and northern North Sea in the Norwegian Deep)

Input data

A preparatory data evaluation workshop was held October 18-22. Though the reviewers were not present for this workshop, on the basis of the summary presentation at WKPRAWN it achieved its stated goals of data preparation in advance of the benchmark.

Commercial data

Quarterly length composition data are available for retained and discarded catch from various years beginning in the earliest in 1990 (SWE 3a retained catch) and the latest in 2019 (NOR 4a discards). Most of the commercial data are combined sex data, except SWE 3a retained, which were sexed up to 2004. While the commercial compositional data forms a patchwork, there are significant overlaps and improving coverage across fleets. It was noted that the Norwegian discard and landings input data 2017-2020 are scarce and considered by the stock experts to be of poor quality (restricted spatial distribution of samples in 2018 in particular). A new sampling scheme is proposed but not yet in place.

Survey indices

The Norwegian shrimp survey is used in the assessment, covering 1985-present with a switch from Q4 to Q2 in 2005 and a switch to Q1 in 2007 when it has been performed since. The length frequency survey data are sex-specific. The method of deriving a total index from the survey was changed from a design-based estimator to a spatiotemporal model that better accounted for the location of stations on strata boundaries and resulted in a more estimates. Given a malfunction of the gear in 2016, the total estimate for that year was considered unreliable (as in previous years). Efforts were made to improve on this (autocorrelated year effects in the index model) but this approach may over-smooth the data for the assessment. The total biomass estimate for 2016 was therefore omitted. Overall, treatment of the survey data was considered appropriate.

Assessment models

Reference model

Stock Synthesis 3 (SS3) is uniquely suited to these multi-fleet, compositional and unbalanced data scenarios. The base model is a two area, two sex, quarterly age-length model that accounts for hermaphroditic reproductive life history of the prawns. A model was fit to include partial historical catches but given the similarity of most parameters, the most recent period were all catches were present was chosen (1970 onwards). An overall growth curve is estimated in the model. Natural mortality in the base model was set based on Chen and Watanabe (1989) using growth parameters from the most recent assessment. Development of a predation index based on the abundance of possible predators for shrimp was explored. While the final index was not used based on a lack of uncertainty measure, I encourage the further development of this approach that has good potential to incorporate ecosystem drivers in natural mortality and can be incorporate in SS3, as was done in one of the runs. Natural mortality is a focus of the ensemble model discussed later. Growth is estimated internally in the model, while maturity and length-weight relationships are inputted, as is common. Selectiveity of the various fleets is modelled as logistic. Recruitment is modelled as a Beverton-Holt function with a suitable value of the recruitment deviation variance set and parameters of the SR relationship estimated within the model.

Most of the parameters of the model were well estimated with the noted exception of the hermaphroditic logit inflection age, which is estimated as close to the lower bound (zero) and well away from the starting value of one. This results in a plausible transition rate but one that has a continually decreasing rate (it is effectively the right-hand side of the ogive). This parameter could be fixed within the model but I would recommend keeping it free in case of future changes.

Five sensitivity runs were compared to the reference run and found not to improve the diagnostics. Though for the reasons outlined in the survey data with respect to 2016, this year was omitted and the "no 2016" run considered the baseline run.

The fit of the reference run to the total survey indices was adequate though some persistent misfit was observed between 1989 and 1996 with the predicted index being consistently larger than the observed index. Fits to the recent time period are good, with the exception of the 2016 data highlighted previously.

Compositional fit to NORSURVEY3a: female data have two peaks in the observed length proportions (1984-2002) but only one estimated in the model. The peaks persist post 2003 though are reduced and, in some years, absent. These missed peaks are also present in NORSURVEY4a. Could it be that there are two age groups there but the growth model blends them? There is some indication of this peak in larger individuals in the catch data also.

The surveys changed from season 4 survey to season 2 survey in 2003 - residuals flipped for both sexes. Survey moved to season 1 in 2006.

NORSURVEY4a post 2003: males underestimated, females overestimated consistently. This is also present in the sex ratios that are consistently biased with the direction of the bias switching pre and post 2003.

On the basis of the residual diagnostics for the base model, it was recommended that for future work these residual patterns should be better understood and corrected, possibly by allowing selected growth parameters to vary, comparison of growth across the sexes, and all with respect to the timing of the survey switches in quarters.

Retrospective analysis showed that estimated the fishing mortality and SSB had relatively low bias but forecasted F was consistently biased high (Mohn's rho 0.35).

A suite of additional diagnostics were provided including run tests, "hindcasting", and MCMC sampling.

Runs tests were used as a measure of goodness of fit by testing multiple aspects of the residuals including a Wald-Wolfowitz runs test, a measure of RMSE and outlier detection. The base case failed in 6 out of 10 of these tests with the argument given that failure is common in these tests. This is a somewhat unsatisfactory outcome in that a test is proposed and then when failed noted that failure is common. It is recommended that measures of goodness-of-fit be further explored in the future, including those that incorporate sex-specific differences in the prawns here and more broadly those that incorporate a measure of model complexity in terms of number of parameters.

Hindcasting, conditions on the catch to forecast observed data and compares the quality of the forecasts against a static assumption. The base model performed well in hindcasting according to the available literature on mean absolute scaled error.

MCMC sampling of the parameters also produced stable results compared to the maximum likelihood solution and demonstrated a robustness and identifiability of the model and parameters. The base model performed relatively well in these developing area of novel diagnostics.

Consideration of this broader suite of diagnostics (and some additional diagnostics such as starting values) is a welcome development and demonstrates the clear intention to improve interrogation of the model fit.

Ensemble model

An ensemble model was developed to incorporate uncertainty in natural mortality. A distribution of possible natural mortalities at age was constructed based on a suite of M life history and other type models. Three M scenarios were developed: a low, median and high M. Rather than average these values and run a single model, the benchmark model ran each of them, quantified their performance on a suite of diagnostics (29) and based on pass/fail criteria provided a weight to the run. This weighting approach falls into the category of "tactical weighting" (Jardim et al., 2021). These final weights were: medium M 86%, low M 69%, and high M 72% (note the weights include compositional data from the survey in 2016 but not the total values).

The joint posterior distribution of the key estimates (assuming a multivariate normal distribution on a suitable parameter scale, e.g., log) was then sampled according to the diagnostic weights. To check the resampling, it was requested that the same M value be combined three times, the resulting distribution showed that this simply resulted in more samples of the estimates and thus behaved correctly in this respect.

Two key questions emerge from these important investigations: 1) why not use the expected value of M from the distribution and fit a single model?; and 2) what are the implications of weighting each diagnostic equally?

While testing three trends of M is not testing a structural uncertainty, it is analogous to testing imperfect initial conditions of hurricanes where multiple forecasts from a single model are combined (different model combinations are also used in hurricane tracking). Given the nonlinear nature of the assessment model, it is not necessarily expected that average of the outputs will be the same as a single model with the average used for M, particularly not in the uncertainty associated, which reflects uncertainty in M.

Equal weighting of diagnostics intuitively assumes that the quality of the fit to the data reflects the plausibility of a given state of nature. The approach suffers, however, from diminishing individual diagnostic weights with increased number of diagnostics; some of which may be more important than others and correlated with others. As a first defensible attempt to combine models on the basis of a wide set of diagnostics, the assessment scientists have done a commendable job.

The decision of whether to accept or reject the ensemble model for me hinged on the logic of the model combination, which appeared well thought-out (M scenarios, diagnostics of runs, delta multivariate log normal samples combined for metric of interest). The procedure is good practice with respect to the published literature while also at the forefront of our understanding of ensembles that have proved their utility in other complex dynamical settings. The assessment scientists addressed the reviewers questions and no plausible reason for rejection could be found that didn't hinge on an argument against novelty. We therefore accepted the ensemble as a basis for advice.

Future recommendations will include a focus on testing the performance of various weighting alternatives, including measures based on model parsimony.

Reference points

The proposed target $F_{MSY} = F_{35\%}$ (F that results in an equilibrium biomass of 35% B0 = B_{target}), is a proxy based on the simulation work of WKREF 1. Use of a proxy F is a departure from the previous approach that was defended on the basis of simulation work that showed that the probability of hitting B_{lim} was unacceptably high for deterministic F_{MSY} (e.g., that obtained from FLBRP in FLR). A combination of $F_{35\%}$ and a trigger point of $B_{trigger} = 0.8\ B_{target}$ was found via short-cut MSE to safeguard the stock with respect to B_{lim}. This was also found to largest yields over the longterm. The differences from deterministic F_{MSY} relate to stochasticity in the form of recruitment process errors that result in the deterministic F_{MSY} being too high. It is important to note that deterministic F_{MSY} is not what would be estimated from eqsim, which has stochastic projections and would therefore suggest a lower F_{MSY} than the deterministic case (Bordet and Rivest, 2014). No eqsim runs were available for comparison.

The limit reference point was initially suggested as $B_{lim} = 10\%$ B_0. Long-term nature of the assessment provides a reasonable estimate of B_0. On inspection, this value lay below historically observed spawning stock biomass. We therefore suggested that the B_{lim} value be raised to 15% B_0 on the basis that it is in the observed range and close to B_{loss} used previously and to where a good recruiting year class has been observed.

In the context of the ensemble model, quantities of interest such as status can be combined from weighting samples from the relative statuses from the individual model for the purpose of advice.

Recommendations for future work

A great deal of novelty was brought in during this benchmark. The approaches were presented in detail with access given to all outputs from the base model. Here I recommend areas for future work to continually improve this assessment and resulting advice.

Improvement in residual pattern with respect to males and females in the survey data. This flipped in 2004. The growth models and transition rates of the hermaphroditic function with respect to season of the survey seems a likely place to work on.

Derive a measure of uncertainty for the predation index to see if it is tracking a trend or is highly uncertain.

Continue with the spatial model to further understand the spatial aspects with respect to prawn in the area from spawning to settlement. From the work conducted for this benchmark, the biomass status of both areas is low but lower in 4a east than 3a. It is exceedingly useful and a credit to the updated assessment to know this.

Diagnostics: It is recommended that measures of goodness-of-fit be further explored including those that incorporate sex-specific differences in the mean lengths. More broadly, explore goodness of fit with measures that incorporate a measure of model complexity in terms of number of parameters. Is the omnibus runs test proposed here better than the individual components of goodness of fit?

Ensemble weighting: explore the performance of various weighting schemes in terms of management performance.

Compare proxy reference point performance with eqsim reference points.

References

Bordet, C., and Rivest, L. P. (2014). A stochastic Pella Tomlinson model and its maximum sustainable yield. Journal of theoretical biology, 360, 46-53.

Jardim, E., Azevedo, M., Brodziak, J., Brooks, E. N., Johnson, K. F., Klibansky, N., Millar, C. P., Minto, C., Mosqueira, I., Nash, R. D.M.,

Vasilakopoulos, P., and Wells, B. K. (2021). Operationalizing ensemble models for scientific advice to fisheries management. ICES Journal of Marine Science, 78(4), 1209-1216.

Northern Shrimp in subareas 1 and 2 (Northeast Arctic)

Input data

A preparatory data evaluation workshop was held October 18-22. Though the reviewers were not present for this workshop, on the basis of the summary presentation at WKPRAWN it achieved its stated goals of data preparation in advance of the benchmark.

Commercial data

A thorough investigation into a CPUE index for the commercial fleet was conducted. This included new and old logbook fishing gear codes as well as strata, vessel size and identification, monthly and spatial effects. Residual diagnostics showed more variability in the tails than expected. This would typically alter the index precision rather than the mean values, which are used in the assessment. Residuals were thoroughly investigated to test for technical creep. A potentially worrying aspect of the survey data coverage is the contraction over time of the commercial fleet with indication of an inshore versus offshore structure to fishing. Further investigations of this and the contraction with respect to the survey data are recommended in future recommendations. Spatial models would not easily converge and had little impact on the index, therefore non-spatial models were chosen for consistency of convergence.

Survey indices

The two main fishery-independent surveys were the joint Norwegian/Russian Barents Sea Ecosystem Survey (BESS) and the Winter Survey. A spatiotemporal modelling approach (sdmTMB) was presented and compared to the design-based estimators currently used. The modelled approach handled changes in coverage well (absence of Russian data in some periods). We requested that the spatiotemporal models be peeled to see if the estimation was consistent, which it was. We therefore accepted this approach as it provides a robust index of abundance for the assessment while also providing information on the spatial distribution of the stock within the survey area and time.

A proposal to combine the BESS and Winter Surveys was not accepted owing to a difference in timing that included the fishery period.

Assessment models

The previous assessment model was a discrete time state space surplus production model implemented in BUGS. The proposed model was SPiCT and comparisons made with this.

Priors were carefully considered, particularly that for carrying capacity K, which is poorly estimated from the available data. A prior was derived from the West Greenland shrimp stock, which was considered to have similar biological properties by the stock experts. Habitat quality was assessed by comparing the mean densities over the period 1988-2002, which found that the mean density of the Barents Sea stock was 31% that of the West Greenland stock. Habitat was thus converted to west Greenland equivalence and multiplier (2.26) for the West Greenland K with the distribution approximated by a lognormal distribution with added uncertainty. The approach to deriving a K value from a similar stock was well thought out, biologically relevant and of benefit to the updated assessment. Future work could consider maximal densities as of use in developing habitat-based priors.

Similar overall trends were found between the original BUGS model and SPiCT. A lag was apparent between the models, possibly reflecting the survey timing or differences in index weighting between the models (lags are apparent between the survey and commercial CPUE indices). Updated indices were the main driver of the change in estimated B/B_{MSY} status while an update in the priors and indices changed the estimated F/F_{MSY} status.

A high correlation was observed between catchability terms and carrying capacity. To explain the high precision of B/Bmsy, assuming a single index (I) and ignoring measurement error, the B/Bmsy in a Schaefer is

$$B/Bmsy = (I/q) / (K/2) = 2I / (qK)$$

The term qK absorbs the negative correlation (high q and low K can give similar value to low q and high K value) so this is why we see narrow confidence limits on the estimated B/Bmsy.

In comparisons with data-rich ICES assessments Bouch et al. (2021) is that when a very high correlation is observed between q and K, the scale of B/Bmsy can be biased so a precise but potentially biased estimate. As highlighted during the benchmark, this reflects the contrast in the data. Based on the estimated production F/Fmsy is poorly estimated as only observations are found on the right side of the limb with broad loops of process variability. But because it is quadratic, the rise is the same as the fall. The scale of both B/Bmsy and F/Fmsy are difficult to estimate for this stock. The prior work on K is therefore key to acceptance of this assessment.

One-step ahead forecast residuals were autocorrelated for the survey indices, highlighting that the model is fitting mostly to the commercial CPUE data. This is also reflected in the measurement error variances of the survey indices being an order of magnitude greater than the commercial CPUE data. Retrospective patterns were acceptable (though the relative F retro was -0.123 but within acceptable intervals for short-lived species).

Multiple sensitivity runs were performed to defend the proposed run. The sensitivity runs including omitting the K prior, commercial CPUE typically showed poor diagnostics relative to the proposed run.

Reference points

The ICES advice rule is implemented within SPiCT using the model-derived F_{MSY} with uncertainty correction and is proposed to be used as the basis for catch advice.

Recommendations for future work

Proposed recommendations include

Further work on the K prior to see what maximal as opposed to average densities can inform on what the limiting habitat could maximally support.

Understanding the differences between the commercial and survey indices. As recommended by Ewen, this could include limiting their derivation (for comparison purposes) to the area of overlap.

Investigate historical data that could improve the estimation of K and thus the high correlation of K and catchability within the model.

Investigation of potential inshore versus offshore commercial CPUE derivation.

Reference

Bouch, P., Minto, C., and Reid, D. G. (2021). Comparative performance of data-poor CMSY and data-moderate SPiCT stock assessment methods when applied to data-rich, real-world stocks. ICES Journal of Marine Science, 78(1), 264-276.

Northern Shrimp on the Flemish Cap (NAFO Div. 3M)

No updated assessment was provided for the Flemish Cap stock though on the basis of the data preparation workshop there appears to be the potential to develop a SPiCT assessment (likely not category 1) for this stock.

Input data

Commercial data

Catches are available from 1993 to 2010 (where there was a fishing moratorium implemented). A standardized CPUE index is available up to the moratorium.

Survey data

An excellent long-term EU survey from 1988-present is available for the stock covering the period of the moratorium. This shows an increase 2016-2019 with a subsequent decrease. Length-composition data are also available from this assessment.

Assessment models

No assessment model was presented owing to a lack of time/resources available to run the assessment.

Recommendations for future work

There is potential to develop a model similar to the advanced 3a4a SS3 model for this stock. That would be technically time consuming requiring much resources. It is therefore recommended to try to fit a SPiCT model to these data. The presence of the moratorium and subsequent increase provides an excellent opportunity to estimate the r parameter. Carrying capacity for this stock is likely to be more challenging but could borrow strength from other stocks, as with the Barents Sea assessment priors. If a SPiCT model were implemented care should be taken with the F process, as it is not a random walk here.

Benchmark workshop on Pandalus stocks (WKPRAWN 2022)

Reviewers' Report: Ewen D. Bell

Northern Shrimp in Divisions 3.a and 4.a East (Skagerrak and Kattegat and northern North Sea in the Norwegian Deep)

Background

This stock was last benchmarked in 2016 with a SS3 model subsequently being used.

This benchmark has addressed issues described by the stock assessors covering both the input data and aspects of the modelling approach.

Input data

The benchmark has made several improvements to the basic data streams. Much of the overhaul of the basic data was covered in the data compilation workshop which did not include reviewers. There remains an issue around the conversion factors used for deriving live-weight estimates for catch boiled at sea but the assessors are awaiting the results of further experiments.

Survey data

The calculation basis of the survey index has been revised and now uses a statistical approach which is considered to be an improvement in robustness over the previous survey-design estimator. Both the biomass index and the length distributions are processed using this new approach.

The 2016 survey suffered from a technical issue with different warp lengths being deployed on either side thus altering the geometry of the net. It was, at the time, considered inappropriate to use the survey catch rates from 2016 within the time series and the Group retained this conclusion. Due to the nature of SS3, the model performs better when the 2016 survey index is left in the series but removed from the likelihood function. The length distribution was considered to be unaffected by the warp length issue and could therefore be used by the model.

Natural mortality.

Natural mortality is now being considered as age-varying (previous assumption was for constant M across ages). The assessors have used the Barefoot Ecologist's toolbox to derive the envelope for M via a range of techniques based on life-history. The 5th, 50th and 80th centiles of the distribution were selected as candidate values for M in the later ensemble method. The 80th centile was chosen over the 95th centile as the distribution has a long upper tail which was considered unrealistic. The use of these age-varying M estimates is considered to be an important improvement.

Use of a time-varying natural mortality where a predation index is used to modify the underlying M was explored in the benchmark. Questions remain as to how best to generate the predation index including the range of species (and size range of predators) that this uses. The application of this approach would represent a significant step towards greater ecosystem considerations and should be pursued in the future.

Model

There are two changes to the modelling approach presented to the benchmark group. The base model remains an SS3 model, but with a number of changes. It is, however, the extension into an ensemble modelling framework that represents a more fundamental change.

SS3 changes

Fleet and area disaggregation: The new model formulation now allows for multiple fleets operating in two areas to better reflect the nature of the fishery and availability of data.

Sex disaggregation: *Pandalus borealis* are protandrous hermaphrodites (change from male to female) and consequently the status of spawning biomass is more appropriately measured as female only. In the previous model this was estimated by using a fixed maturity ogive for females. The new model formulation uses the sex-disaggregated length distributions from the survey and some catch sampling (Swedish retained catch in 3a) to estimate the sex-change ogive (which is assumed to be a function of length).

Growth (Von Bertalanffy) is fitted within the model and assumed to be common for both sexes.

Model fit

Fits to the new survey biomass estimates are generally good although the fit seems relatively weak for the 2013-2015 period in both the 3a and 4a sections.

The new model structure does generate some distinct patterns in the length frequency residuals, mostly in the survey fits but also in some of the commercial catch series. There is a dominance of positive residuals for females in the period up to around 2002 and a dominance of negative residuals for females in the latter period. The converse is broadly true for males. The timing of this change in residuals is commensurate with the change in survey timing as the summer survey ran until 2002, was in autumn 2004 and 2005 and has been in the winter since. The cause of this residual pattern is unresolved and multiple hypotheses present themselves. A seasonal aspect to sex-change (in addition to length) could explain the inflexion of patterns being similar to the change in survey seasonality. Another hypothesis is that growth rate and/or the sex change ogive has been changing over the time period. Both of these have the potential to affect productivity estimates and therefore it is recommended that exploration of the residual patterns, their potential causes and the resulting impacts on the assessment are pursued for the next benchmark.

Model fitting was explored with a range of diagnostic tools including runs tests on the survey index and mean length residuals, retrospective runs and hindcasting. This approach to appraising the model fit is more comprehensive than most assessments and should be explored more widely in the ICES community. The choice of which model diagnostics to generate and compare is potentially critical. It was notable that although one of the key changes to the previous SS3 model was the inclusion of sex-disaggregation, none of the diagnostics included sex-disaggregated elements, instead focussing on broader features (e.g. mean length). Future modifications to the diagnostic tests could explore making more of these specific model features. Care should be made to ensure that the inclusion of relatively insensitive and/or correlated diagnostics does not result in a biased interpretation of overall model performance. As is discussed below, the use and weighting of multiple diagnostics remains a topic of active research.

Overall the diagnostic tests indicate that the model is performing sufficiently well and is appropriate for this stock.

Ensemble model

The majority of stock assessments pick a single model as the “best” (or possibly “least worst”) description of the system. This approach ignores the often considerable uncertainty around model structure and/or key parameter assumptions. The benchmark group was presented with an ensemble modelling approach to address the uncertainty around the key parameter of natural mortality. Three SS3 model runs were generated (low, medium and high natural mortality) and the suite of model diagnostics generated. For each diagnostic a hard pass/fail threshold was set and then the proportion of passes was used to provide a weighting for each model. This particular application assumes that each of the 29 diagnostics used is considered equally important. The thresholds for the pass/fail status of each diagnostic as well as the number and weighting of the diagnostics have the potential to be highly influential on the final result of the ensemble model. In the future, consideration should be given to whether there is correlation or independence between the diagnostics, the number of diagnostics chosen and their relative weighting. The model developers did say that diagnostic weighting was one of the hot topics for the emergent world of ensemble modelling and I would expect that future model development will embrace any new "best practice".

The ensemble model approach presented to the group represents a new phase for stock assessment which is able to incorporate a greater level of uncertainty than the more traditional "pick your best model" approach. Although the methodology is still developing and maturing it offers tangible benefits to stock assessment and the approach presented is therefore considered an appropriate basis for the generation of advice in this case.

Reference points

Moving to an ensemble model does require a shift in the concept of reference points as the competing models must be treated in a relative sense. In this particular application the different assumptions of natural mortality result in substantially different absolute levels of productivity but comparison can be made in relative terms to the virgin biomass in each scenario. This makes the reference points more akin to proxies than direct estimates but this does not detract from their appropriateness. The determination of reference points was established using a shortcut MSE approach, testing a range of fishing mortalities and biomass thresholds in relation to virgin stock size (B_0). Fishing mortalities that resulted in biomasses of 30%, 35% and 40% of B0 were explored, this range being well established in global fisheries. A range of multipliers (100%, 80% and 60%) on B_{target} (where B_{target} was 30% 35% and 40% of B_0) was explored as candidates for $B_{trigger}$. In the initial simulation work B_{lim} was defined as 10% of B_0 however the Group subsequently considered this to be too low as it was outside the range of observed biomasses and recovery from the very lowest biomasses had been slow. The Group agreed that B_{lim} for this stock should be 15% of B_0 as that placed it within the observed range where reasonable recruitment has been generated and was reasonably analogous to the use of B_{loss} in other deterministic stocks.

The combination of an F_{msy} proxy at $F_{30\%}$, combined with $B_{threshold}$ at 80% of $B_{30\%}$ satisfied the criterion of being above B_{lim} with >95% probability and generating catches within 95% of MSY. Interannual variation in catch was not a consideration in the selection of reference points.

One drawback with this approach to estimating reference points is the time taken to run the MSE which places constraints on the granularity of the search grid (9 combinations of F and B). Given the advances in modelling approach taken for this stock it would be preferable to have a finer grid of combinations to maximise the potential of the endeavour.

Northern Shrimp in subareas 1 and 2 (Northeast Arctic)

Exploration of Ecosystem drivers

The assessors have spent considerable time exploring the possibility of including ecosystem drivers into this assessment which is commendable. In the end neither of the hypothesised drivers seemed to have a significant relationship with *Pandalus* dynamics however these types of explorations should continue (and not just for this stock!).

Input data

Survey

Two surveys are available for this stock, the Barents Sea ecosystem survey and the Barents Sea demersal fish survey. The ecosystem survey operates in the summer and has broad spatial coverage however it has been plagued by technical difficulties in the last few years. The Barents Sea demersal fish survey operates in the winter and is spatially constricted due to ice cover but has been operational in years when the ecosystem survey has encountered problems. The benchmark group explored the potential to use a combined index from both the ecosystem survey and the winter fish survey to provide a more robust estimate in years when the ecosystem survey was incomplete however this was rejected due to some convergence issues and the fact that a large portion of fishing occurs between the two surveys.

The winter survey, although there were some significant correlations with one strata of the ecosystem survey, was considered to spatially limited to offer an index of the whole stock. The Group therefore concluded that the ecosystem survey should form the basis of the survey index. The construction of the index has been changed from a survey-design based approach to a statistical approach that factors in depth, time of day and spatial correlation. This new index is considered to be a more robust estimate of stock abundance than the previous approach.

CPUE

Commercial CPUE is used within the model and also underwent a comprehensive overhaul. The basic input data were reanalysed to address issues pertaining to changes to Norwegian logbook formats. The generation of the standardised CPUE index was also revised to a mixed-effects model approach which is considered more robust than the previous approach. A common concern with the use of commercial CPUE is that of technical creep and the assessors have concluded that no residual technical creep is evident. There remains some spatial patterning of residuals that the assessors are aware of.

Contraction of the spatial coverage of the commercial fishery is evident, which means there is a mismatch between the portions of stock covered by the survey and fishery-dependent information. Potentially causes a tension in the system by using both data pertaining to the wider stock and data pertaining to the fished stock. This may be a reason for the apparent difference in time series trajectory between the survey biomass and commercial CPUE trends. Further exploration of this issue is recommended for the next benchmark.

Assessment model

This stock is assessed using surplus production techniques. Collection of length-based data has been patchy and insufficient for a length based assessment. The slower growth rate of this stock also means that there are few if any modes in the length distributions for models to effectively determine year classes.

The previous assessment model was constructed in the Bayesian OpenBugs framework however this was incompatible with ICES TAF concept and so the decision had been made to move to the SPiCT framework that is now in common use within ICES. Although some of the features and flexibilities of the OpenBugs framework were not available within SPiCT the final model choice appears to offer very similar outcomes (even if the underlying philosophy is markedly different).

A detailed explanation was given for the choice of priors for key parameters with a particular focus on the prior for carrying capacity (k). This is poorly estimated with data from this particular stock, so the assessors explored the potential to use information from another *Pandalus* stock (West Greenland). The approach presented appears reasonable but is reliant upon the assumption of equal catchability for the two trawl surveys and a similar level of exploitation between the two areas (due to the use of mean catch rates). An alternative to the use of mean catch rates could see maximum catch rates used as these may be closer to the unexploited state and it is recommended to explore this for the next benchmark.

Model fits

The move from the Bayesian model to the SPiCT framework makes relatively little difference to the overall trend in mortality and stock size although for some reason the SPiCT assessment does seem to lag one year behind the Bayesian model. This was noted in the workshop and it is hypothesised that the difference in survey vs CPUE trends may be partially responsible but further investigation on this would be welcomed for the next benchmark.

A range of candidate model configurations were presented (including removal of the k-prior and removal of the commercial CPUE) and whilst the reference model did not satisfy all criteria it was considered to perform satisfactorily.

Northern Shrimp on the Flemish Cap (NAFO Div. 3M)

No new work was presented to the benchmark workshop.

7 Recommendations

The following recommendations were made by WKPRAWN:

No	Description	To whom
1	To improve understanding a flow chart of various alternative model setup investigated during a benchmark process should be included in the guidelines for benchmarks.	BOG
2	WKPRAWN recommends that conversion factors (the correction factors for weight loss from boiling+cooling and for weight gain from storing raw shrimp on ice) and raising procedures are investigated thoroughly before the next benchmark and this work should be coordinated with WGCATCH and involve the relevant *Pandalus* experts	WGCATCH & NIPAG
3	Differences in growth between the Skagerrak and the Norwegian Deep as well as sex differences in growth should be further researched. The assessment diagnostic show strong patterns in the residual fits to the length distributions. WKPRAWN suspected that these are most likely due to changes in growth or difference in growth across time and area not accounted for in the way growth is currently modelled	NIPAG
4	It is strongly recommended that issues concerning stock structure of pra.27.12 be further investigated and considered in the future	NIPAG
5	Recommended at regular intervals to revisit the Length frequency data for pra.27.12 to ensure that data quality is sufficient to potentially found the estimation of recruitment indices.	NIPAG

Annex 1: List of participants

Name	Institute	Email
Alessandro Orio	Swedish University of Agricultural Sciences, Sweden	alessandro.orio@slu.se
Anndorte Burmeister	Greenland Institute of Natural Resources	anndorte@natur.gl
Carsten Hvingel	Institute of Marine Research, Norway	carsten.hvingel@hi.no
Chris Griffiths	SLU Department of Aquatic Resources Institute of Marine Research, Sweden	christopher.griffiths@slu.se
Cóilín Minto	Galway-Mayo Institute of Technology, Ireland	Coilin.Minto@gmit.ie
Colm Lordan	Marine Institute, Ireland	colm.lordan@marine.ie
Ewen D. Bell	Cefas Lowestoft Laboratory, UK	ewen.bell@cefas.co.uk
Fabian Zimmermann	Institute of Marine Research Tromsø, Norway	fabian.zimmermann@hi.no
Francesco Masnadi	University of Bologna, Italy	francesco.masnadi@studio.unibo.it
Guldborg Søvik	Institute of Marine Research, Norway	guldborg.soevik@hi.no
Johan Lövgren	SLU Department of Aquatic Resources-SLU Aqua, Sweden	johan.lovgren@slu.se
Katherine Skanes	Fisheries and Oceans Canada	Katherine.Skanes@dfo-mpo.gc.ca
Katja Norén	SLU Department of Aquatic Resources-SLU Aqua, Sweden	katja.noren@slu.se
Krista Baker	Northwest Atlantic Fisheries Center, Canada	krista.baker@dfo-mpo.gc.ca
Mark Simpson	Fisheries and Oceans Canada	mark.simpson2@dfo-mpo.gc.ca
Massimiliano Cardinale	SLU Department of Aquatic Resources-SLU Aqua, Sweden	massimiliano.cardinale@slu.se

Name	Institute	Email
Michael Wall Andersen	Danish Fishermen Producer Organisation_Fredericia, Denmark	ma@dkfisk.dk
Mikaela Bergenius Nord	SLU Department of Aquatic Resources-SLU Aqua, Sweden	mikaela.bergenius.nord@slu.se
José Miguel Casas Sanchez	The Spanish Institute of Oceanography Centro Oceanográfico de Vigo	mikel.casas@ieo.es
Ole Ritzau Eigaard	DTU Aqua, National Institute of Aquatic Resources, Denmark	ore@aqua.dtu.dk
Patrik Börjesson	SLU Department of Aquatic Resources-SLU Aqua, Sweden	patrik.borjesson@slu.se
Rui Catarino	International Council for the Exploration of the Sea	rui.catarino@ices.dk
Sergey Bakanev	Polar Branch of Russian Research Institute of Marine Fisheries and Oceanography, Russian Federation	bakanev@pinro.ru
Tanja Baagoe Buch	Greenland Institute of Natural Resources	TaBb@natur.gl

Annex 2: Working documents

WD 1: Historic landings of northern shrimp (*Pandalus borealis*) in Norway. Guldborg Søvik.

Introduction

The Norwegian fishery for northern shrimp (Pandalus borealis), hereafter shrimp, started by mere chance in the late 1890s, as the large stocks of shrimp in the Oslofjord and eastern Skagerrak region were discovered through fishery investigations with other purposes than finding shrimp (Hjort and Ruud 1938). In 1887, Johan Hjort discovered high densities of shrimp in the Langesundsfjord in southern Norway. Before 1887, large, harvestable stocks of shrimp were only known from the Drammensfjord, where a shrimp fishery by hand net had taken place.

A profitable shrimp fishery was established only a year after the discovery of the new resource. In 1899, 10 vessels participated in the shrimp fishery in the Langesundsfjord, and in 1901, 11 vessels trawled for shrimp in the inner parts of the Oslofjord.

Much work was devoted to the development of suitable gear, equipment of vessels, mapping of shrimp fields, and development of a new market, mainly driven by the fishers themselves. The fishery was first conducted by sailing vessels, but these were rather quickly replaced by steam and motor boats. During World War I sails were again utilized due to fuel shortages. Shrimp trawling with sailing boats in the small and narrow shrimp fields along the Norwegian coast was challenging. As new shrimp fields were discovered along the coast, the fishery gradually moved westwards.

Official landings statistics are available only from 1908 in "Norges Fiskerier" (Norwegian fisheries) (Hjort and Ruud 1938). This year, 405,8 tonnes of shrimp were landed, and the shrimp fishery was by then conducted along the coast from the Østfold county to the Rogaland county.

Since 1908, landings statistics by year and county, and for many years also by municipality, are available through "Norges Fiskerier", as scanned reports. The data have not earlier been digitized, and as such, have not been readily available. An overview over the Norwegian historic shrimp landings is an important part of the history of the Norwegian fisheries. Landings statistics are also important input data to stock assessment models. The goal of our work is threefold: 1) provide the first description of the Norwegian shrimp landings by county from the infancy of the fishery until today, 2) provide an equivalent description for the counties from Hordaland to Nordland, a region where the shrimp fishery can be characterized as being a purely coastal and fjord fishery, and finally 3) provide historic Norwegian landings by statistical area as input data to the current stock assessment model for the shrimp stock in the Skagerrak and Norwegian Deep (pra.27.3a4a).

Material and Methods

Data of shrimp catch and value is collected from the SSB landing register, for the time period of 1908-1978, from 6 Fiske and 10.05 Fiske og fiskeoppdrett.

The data is punched in Excel from SSB fishery reports. Table 4 is used in year 1908-1952, table 5 for 1953-1962, table 1 for 1963-1975, table 2 for 1976 and table 20 and 21 for 1977-1978. Counties are updated to current place.

Values for 1977-2021 are collected from a digital landing register.

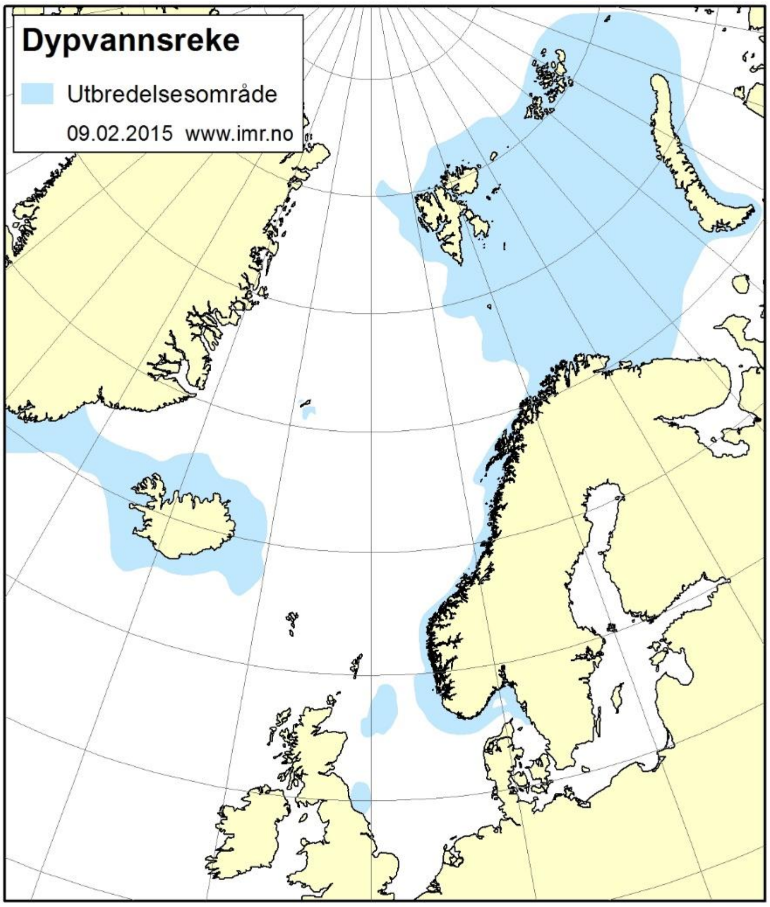


Figure 2.1: Distribution of northern shrimp, Fiskeridirektoratet (2015)

Results

Data collected from SSB landing register, 1908-1978.

Year	County	Place	Landings (ton)	Value (1000 NOK)	Table nr
1908	Oslo		15	9	4
1908	Østfold	Skjeberg	0.1	0.05	4
1908	Østfold	Hvaler	105.475	50.628	4
1908	Østfold	Onsø	100	35	4
1908	Akershus	Nesodden	1.8	0.72	4
1908	Akershus	Aker	0.2	0.1	4
1908	Buskerud	Røken	1	0.6	4
1908	Vestfold	Holmestrand	0.9	0.45	4
1908	Vestfold	Sande	7	3.5	4
1908	Vestfold	Tjømø	60	24	4

1–10 of 3386 rows Previous 1 2 3 4 5 ... 339 Next

Figure 3.1: Data of Norwegian shrimp landings, weight in tons and value in 1000 NOK, 1908-1977

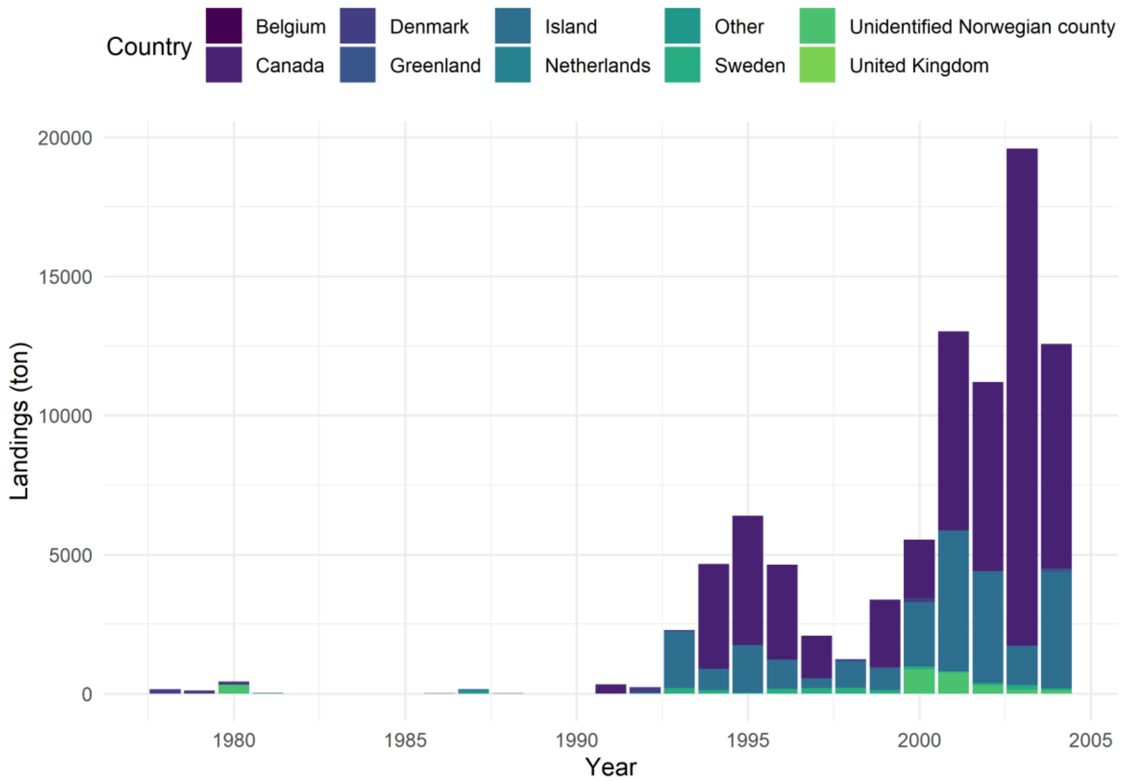


Figure 3.2: Landings registered overseas, for timeperiod 1977-2004.

Year	County	Place	Landings (ton)
1908	Akershus	Aker	0.2
1908	Akershus	Nesodden	1.8
1908	Aust-Agder	Risør og Søndeled	1
1908	Buskerud	Røken	1
1908	Nordland	Saltdalen	0.05
1908	Oslo		15
1908	Rogaland	Hetland	6
1908	Rogaland	Stavanger	82.5
1908	Sogn og Fjordane	Gulen	8
1908	Telemark	Brevik	6

1–10 of 6906 rows   Previous 1 2 3 4 5 ... 691 Next

Figure 3.3: Combined data of total landings per county and place, for timeperiod 1908-2021.

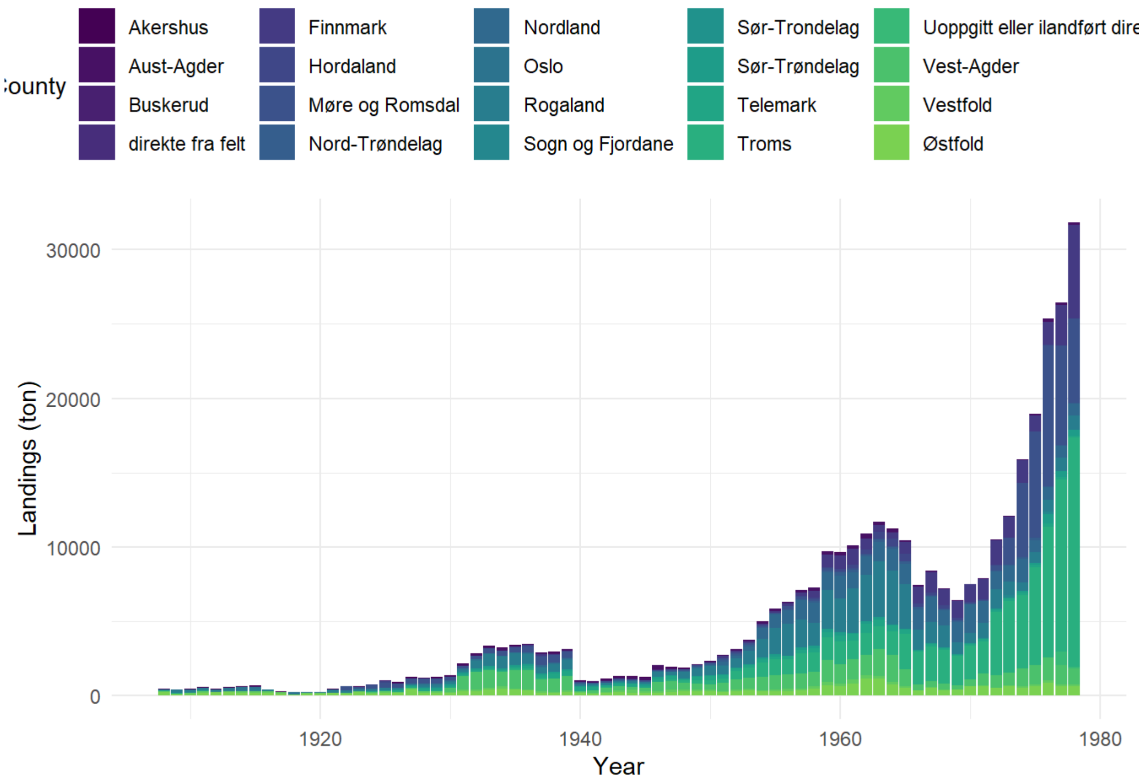


Figure 3.4: Total landings per county, in tonnes, year 1908-1977

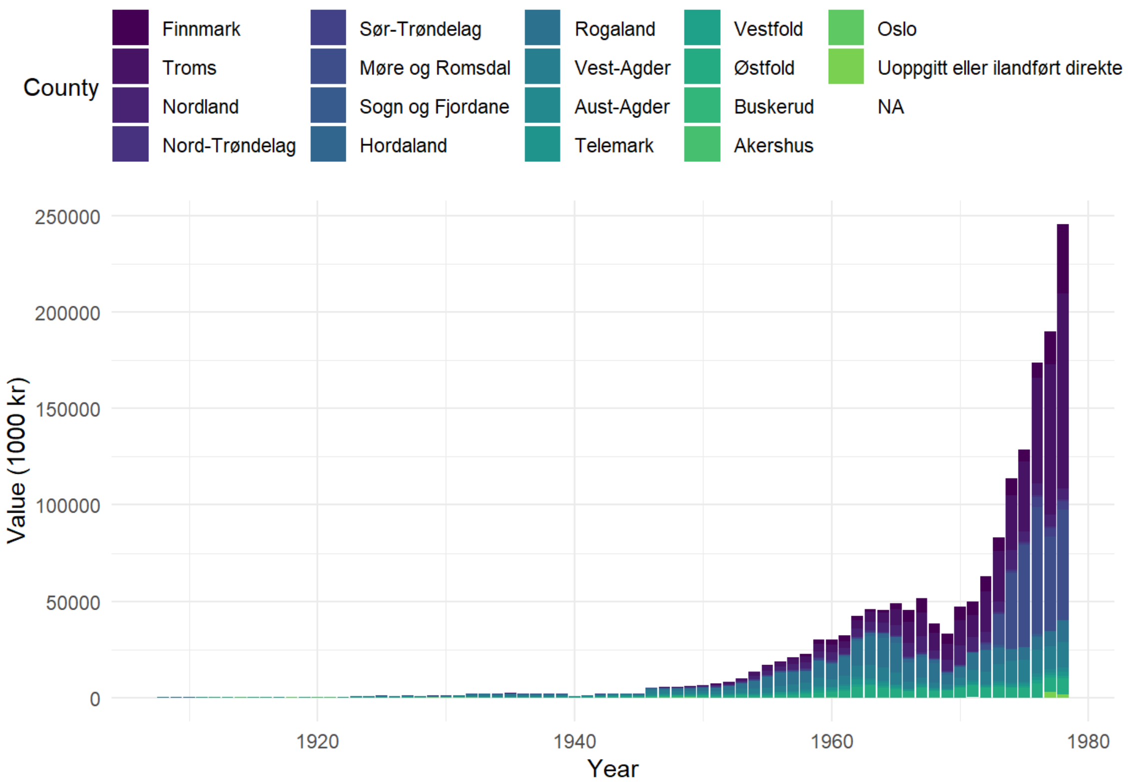


Figure 3.5: Value of landings per county, in 1000 NOK, year 1908-1977

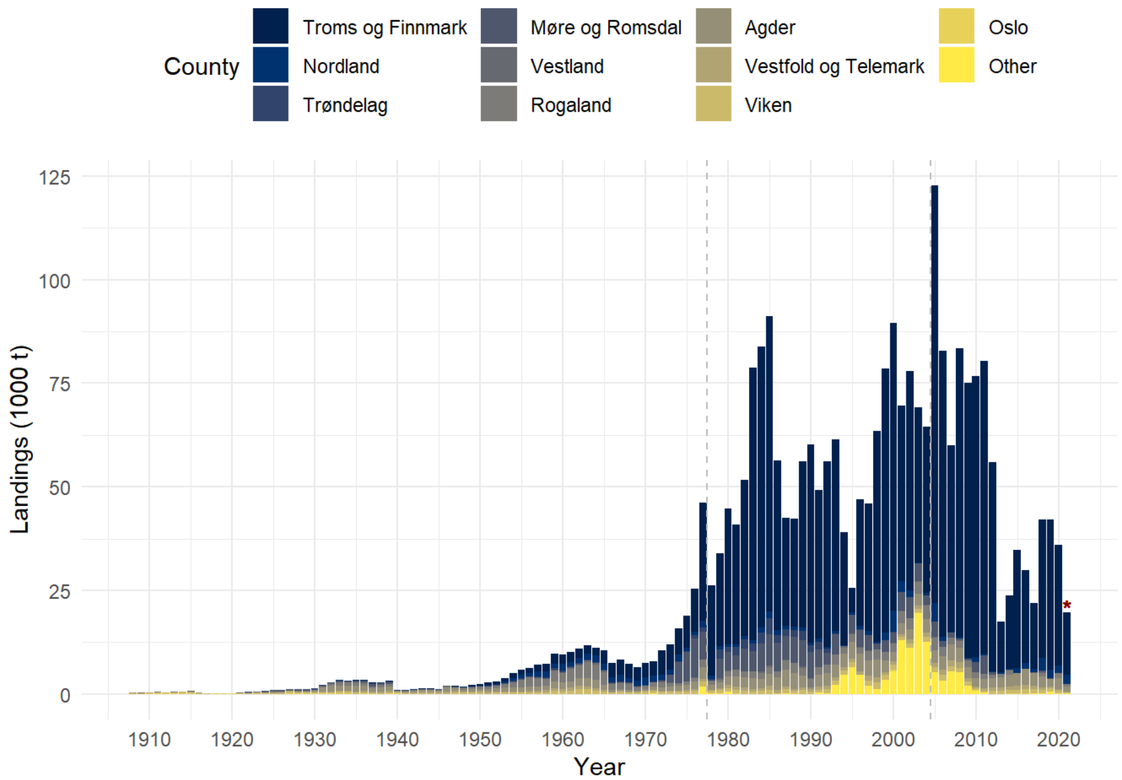


Figure 3.6: Total landings over time per county. Other includes landings abroad or at unknown locations.

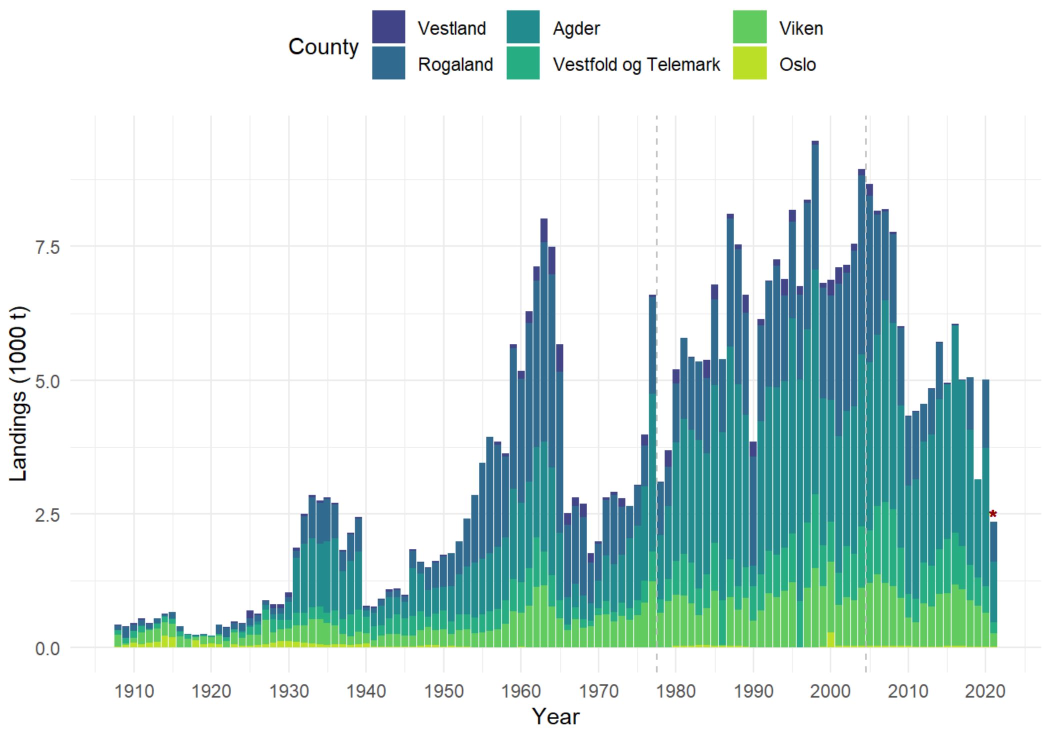


Figure 3.7: Total landings over time in southwestern Norway per county.

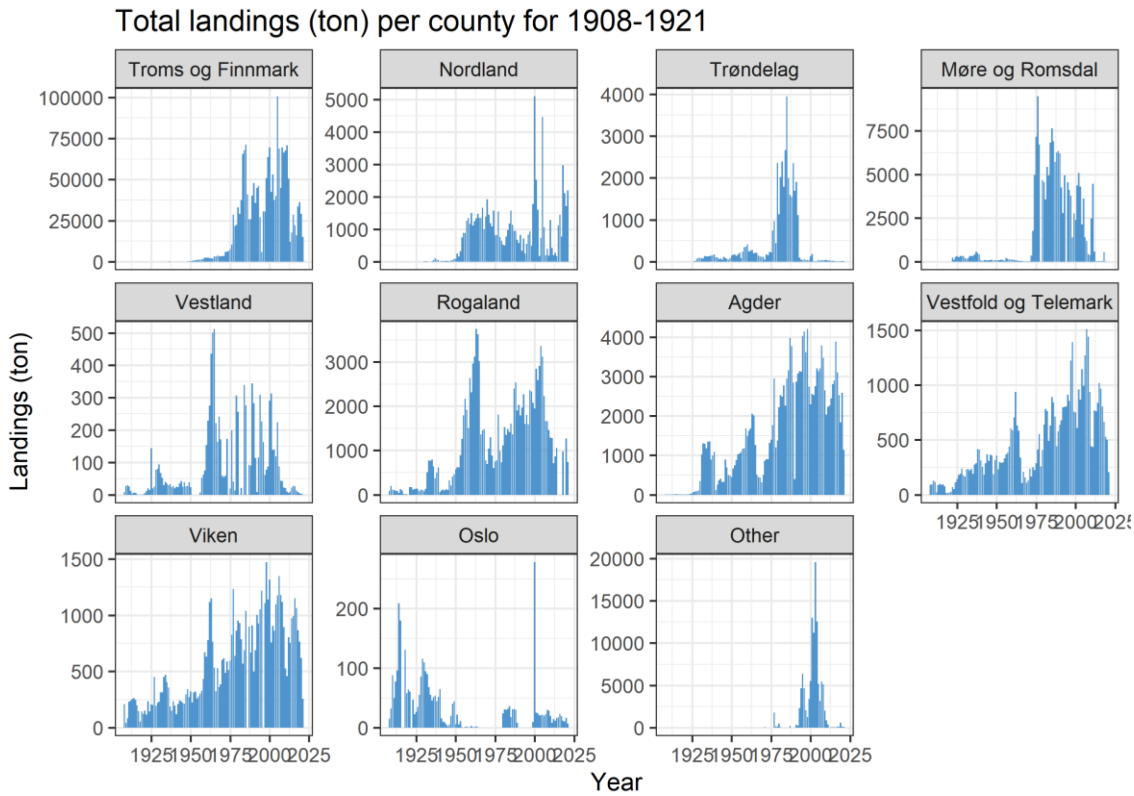


Figure 3.8: Total landings per county, for 1908-2021 (ton)

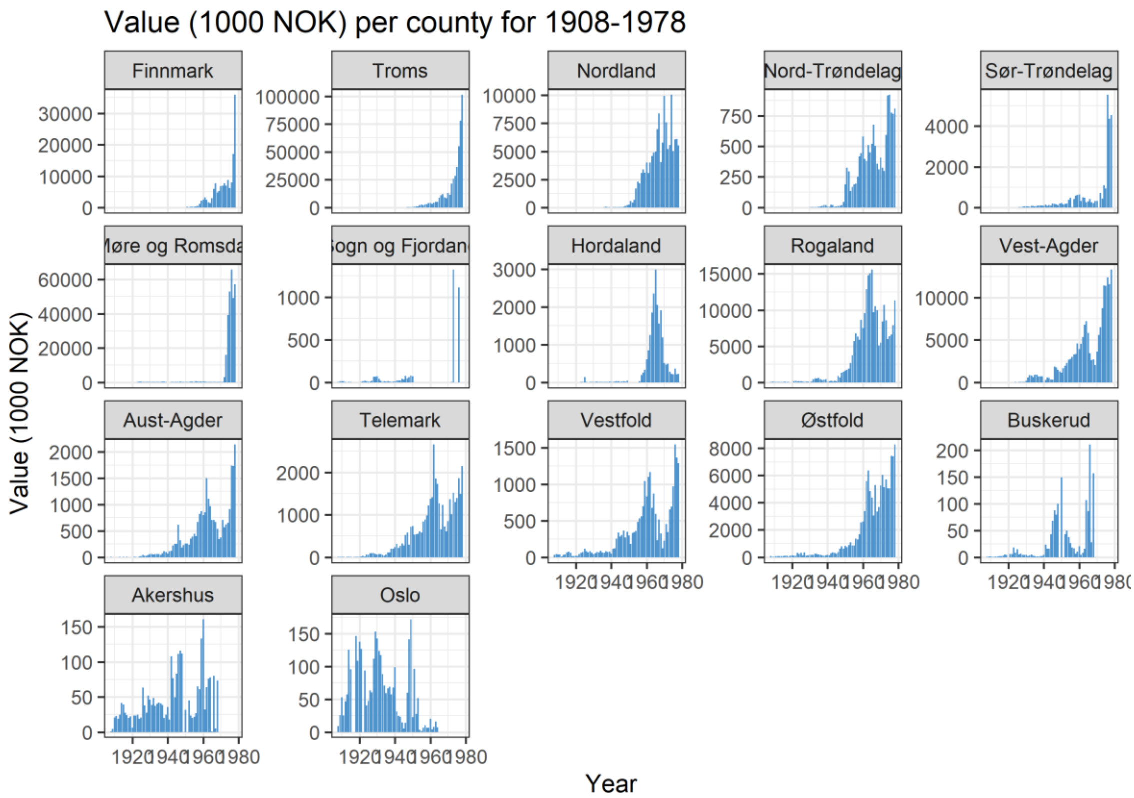


Figure 3.9: Value of landings per county, for 1908-1977 (in 1000 NOK)

WD 2: Fisheries data from the Norwegian commercial northern shrimp (Pandalus borealis) fishery in Skagerrak and the Norwegian Deep (ICES Divisions 3.a and 4.a East)

Guldborg Søvik, Trude H. Thangstad, Fabian Zimmermann

Introduction

This working document describes the Norwegian catch data delivered to the 2021-2022 benchmark of the northern shrimp stock (*Pandalus borealis*) in Skagerrak and the Norwegian Deep (pra.27.3a4a). The Norwegian sampling program has since 2005 sampled unsorted catches from commercial shrimp trawlers, to provide Norwegian catches-per-length as input data to the assessment model. The 2021-2022 benchmark requested length data from discards and landings separately, in order to model discards back in time.

Before 2005, it is unknow if and eventually how Norwegian shrimp catches from Skagerrak and the Norwegian Deep were sampled. No effort has been put into finding out more about the earlier Norwegian catch sampling.

This working document provides a description of the Norwegian shrimp fleet and fishing in Skagerrak and the Norwegian Deep, as well as the two different Norwegian sampling programs, and evaluates how well the former and present sampling program covers the fishery in time and space.

Methods

Statistical grids

Norwegian official landings are registered by statistical area and location (Figure 1a). The Norwegian statistical areas correspond with ICES Divisions (Figure 1b), with some exceptions. Both systems changed the demarcation between Skagerrak and the Norwegian Deep in 2018. The Norwegian areas 08 and 28 correspond to ICES Div. 4.a. East. The Norwegian area 09 correspond to ICES Div. 3.a. The minor Norwegian shrimp landings in some years from the northeastern corner of area 41 (i.e. ICES Div. 4.b) are assigned to area 08 (Div. 4.a).

Norwegian shrimp fishery

Norwegian fleet

The composition of the Norwegian shrimp fleet is shown by number of vessels per length group. Since 2005, this information has been obtained from the landings statistics. Similar information back to 1995 is also available to the Norwegian Institute of Marine Research (IMR).

Fishery statistics

Data on official Norwegian shrimp landings per year, quarter and area were obtained from the Norwegian Directorate of Fisheries, where information on landings per vessel trip, with vessel length information has been available since 2005.

Electronic logbooks with data per haul have been available since 2011. In Skagerrak (area 09) and in the Norwegian Deep (areas 08+28), all vessels respectively ≥ 12 m and ≥ 15 m are obliged to fill in logbooks. In Skagerrak, this pertains to fishing outside 4 nm.

In the shrimp fishery in the Skagerrak and Norwegian Deep, fishers sort the shrimp catch onboard in normally three size fractions (sometimes two): large shrimp, medium-sized shrimp and small juvenile shrimp. The third size fraction may also contain glass shrimp (*Pasiphaea* sp.) and is usually discarded, but may be landed. The medium-sized shrimp are landed raw, fetching a low prize per kilo, while the catch fraction consisting of the large females is boiled onboard to be landed fresh, fetching high prizes (ICES 2021). As shrimp lose weight when boiled, the ICES shrimp working group (NIPAG) corrects the official weight of the boiled landings by multiplying with a factor of 1.13, to obtain live weight. Information on conservation of the Norwegian landings (boiled, raw) has been available from the Norwegian Directorate of Fisheries since 2000.

Discards

Norwegian discards of shrimp have been estimated since 2009. Discards were recalculated for the benchmark in 2022, and were only estimated for areas and quarters with data available. For the years 2009-2016, quarterly discards were estimated by applying the Danish discard ratio per quarter to Norwegian landings corrected for loss of weight due to boiling, assuming the same discard practice in the two countries.

$$Norwegian\ discards\ (Q) = \frac{r * Norwegian\ landings\ (Q)}{(1-r)} \qquad (Eq.\ 1)$$

Here, r is the Danish discard rate (discards as proportion of total catches).

The Danish numbers on landings (corrected for boiling) and discards were obtained from Danish researchers (Table 1). The Danish shrimp onboard observer program takes place primarily in Skagerrak. Data on Danish discards in 2009-2016 therefore exist only from this area, and Norwegian discards from the same years could only be estimated for Skagerrak.

Since 2017, discards have been estimated using data from the Norwegian Coastal Reference fleet (CRF) (Hatlebrekke et al. 2021). The CRF consists of commercial vessels engaged by IMR to register and report all catches including discarded non-commercial bycatch species and discarded fish/shrimp juveniles below minimum legal size (MLS). Since 2016, shrimp trawlers in the Skagerrak and Norwegian Deep have been included in the CRF. For all hauls, the vessels report weight of the three catch fractions (large, medium-sized and juveniles/glass shrimp), and note whether the shrimp are landed or discarded. As the vessels are not equipped with scales, weight of the discarded shrimp is estimated onboard (by using a basket holding a known weight of shrimp). The boiled landings are weighed in landing facilities, while the weight of the landed raw shrimp is estimated as onboard, with a basket. The CRF catch weights of large shrimp have not been corrected for boiling.

According to the instruction, from one haul every second week, vessels send samples from the three shrimp catch fractions (large (2 kg), medium-sized (1 kg), discards (1 kg)) to IMR where the samples are weighed and length measured (up to 100 specimens) (carapace length (CL) in mm). The samples of large shrimp are taken before boiling. The samples from the discard fraction are sorted by IMR into northern shrimp and glass shrimp which are weighed separately. The proportion of northern shrimp in the sample is used for correcting the weight of the shrimp discard in the corresponding haul, in the data from the CRF vessels, such that the total discard weight is partitioned into separate weights for respectively northern shrimp and glass shrimp.

Mean proportion p of northern shrimp in the discard fraction is estimated from the discard samples sent to IMR, per quarter. The discard rate $d.r.$ for sampled hauls is calculated as

$$d.r. = \frac{weight\ (discarded\ Pandalus)}{weight\ (discarded\ Pandalus + landed\ Pandalus)} \qquad \text{(Eq. 2)}$$

The discard rate for un-sampled hauls is calculated as

$$d.r. = \frac{weight\ (discarded\ shrimp * p)}{weight\ (discarded\ shrimp * p + landed\ Pandalus)} \qquad \text{(Eq. 3)}$$

Quarterly discard rate is estimated as a weighted average over all haul-wise discard rates. Due to few vessels in the CRF, an uneven distribution of hauls over statistical areas per quarter and year, and a limited number of trips and samples, discard rates were estimated for areas 09 and 08 combined for the years 2019-2020 (no CRF vessels in area 28). In 2016-2018, there were no CRF vessels fishing in area 08, so discard rates were not estimated for this area these years. Quarterly discards were estimated using Eq. 1, with r as the Norwegian discard rate.

Catch sampling

Since 2005, IMR has paid commercial shrimp fishers for self-sampling their catches. Samples (1-2 kg) are taken from the unsorted catch, before any sorting has taken place, representing the size distribution of the stock available to the trawl. The fishers are asked to take a sample once every month, but some sample less, and some also sample more than once a month. The frozen samples are sent to IMR where the shrimp are sorted to sex and maturity stage and length measured (CL in whole mm). In 2016-2020, samples from unsorted catch were also obtained from vessels in the CRF. The weight of the shrimp catches from which samples are taken, have been provided only since 2020.

Since 2005 with the exceptions of 2008 and 2009, the Norwegian Coast Guard has, as part of their inspections of fisheries, collected shrimp samples from the fishery taking place in the Norwegian waters of Skagerrak and the Norwegian Deep, mainly from Norwegian vessels, but also to a lesser extent from Danish and Swedish trawlers. These samples are also taken from the unsorted catch. The frozen samples are delivered at IMR's facilities where the shrimp are sorted to sex and stage and length measured (CL).

Catch sampling of the CRF catches sorted into three catch fractions, is described above.

Norwegian shrimp catches by length

Since 2006, Norwegian shrimp catches by length have been estimated. Since 2016, these data form part of the input data to the age-based assessment model (Stock Synthesis) (ICES 2016).

Numbers per length (whole mm length groups) are summed for all samples per year, quarter and area.

Numbers per length in the pooled samples (per year, quarter and area) are raised to numbers per length in total catches, i.e. corrected landings + estimated discards (per year, quarter and area), by assuming that the relation between the numbers per length in the samples and in the total catch is the same as the relation between the summed weight of the samples and the weight of the total catch:

$$\frac{No\ per\ length\ (samples)}{No\ per\ length\ (catch)} = \frac{W\ (samples)}{W\ (catch)} \qquad \textit{(Eq. 4)}$$

Here, *No* is numbers and *W* is weight. Thus, for each length bin, the numbers per length in the catch is equal to:

$$No\ per\ length\ (catch) = \frac{No\ per\ length\ (samples) * W\ (catch)}{W\ (samples)} \qquad \textit{(Eq. 5)}$$

Results

The Norwegian shrimp fishery

The shrimp fleet

The number of vessels in the Norwegian shrimp fleet in the Skagerrak and Norwegian Deep has decreased from around 400 in 1995 to presently around 200 (ICES 2021). The fleet structure differs substantially between Skagerrak and the Norwegian Deep (Figure 2, Table 2). The Skagerrak fleet is larger than the fleet farther west and is totally dominated by vessels < 15 m. Since 2005, the only notable change in this fleet is the decrease in the number of vessels 11-14.99 m. In the Norwegian Deep, the fleet segment < 11 m is the most numerous, but the fleet is more varied with also many large vessels. Since 2005, the number of trawlers in all size categories has decreased except for the largest ones ≥ 28 m, which shows a slight increase. The number of medium sized vessels (15-20.99 m) has shown the largest decrease.

Landings

Spatial distribution and temporal trends

The Norwegian shrimp fleet operates along the Norwegian coast, from Oslofjorden to Karmøy (Figure 3). In some years, shrimp trawling has taken place west of Bømlo as well. The fleet also operates along the southern part of the Norwegian Trench, mainly from 5 to 8 ºE. Since 2018, Norwegian vessels also trawl in the Swedish zone. Before 2013, trawling also took place farther north in the Norwegian Deep, but the shrimp density in this area has decreased lately (Søvik and Thangstad 2021). The larger vessels tend to fish offshore (based on the distribution of hauls with twin trawl), while the smaller vessels fish along the Norwegian coast and in the inner Skagerrak (based on the use of single trawl) (Figure 3).

In most years since 1988, landings have been larger in Skagerrak than in the Norwegian Deep (Table 3).

The fishery takes place the whole year (Figure 4, Table 3), but in the Norwegian Deep, landings decline somewhat in the fourth quarter which may reflect periods of rough weather in autumn/winter hindering fishing. In both areas, landings decline in May-June, which could be due to the shrimp disappearing from shrimp grounds in spring after the females have hatched the roe.

Landings by vessel size

In Skagerrak, 50-60 % of the shrimp catches are landed by vessels < 15 m in length, this pattern has been quite stable since 2005 (Figure 5, Table 4). In the Norwegian Deep, only 10-20 % of the landings come from vessels in this size category. Here, vessels 21-26.99 m in length have the largest portion of the landings, and this portion has increased from 2005 to 2020. Landings from vessels 15-20.99 m have decreased in the same time period, from 1302 tons in 2005 to only 12 tons in 2020 (Table 4).

Discards

Discards are expected to be highest in quarters 1 and 2, as the recruiting year class then is below MLS. This is seen for some years, but not all (Table 5). The highest discards were estimated for 2014, and especially for quarter 1 this year. This reflects the large amount of juveniles this year due to the very large 2013-year class (Søvik and Thangstad 2021). This year class also shows up as large discards in quarter 4 in 2013.

Not all shrimp below MLS are discarded. Numbers from the sales organization Fiskehav SA shows that between 2 and 111 tons of shrimp below MLS were landed annually in the years 2016-2021 (Table 6).

Catch sampling from unsorted catches

The number of fishermen (vessels) self-sampling their catches, has varied between years, from only two in 2005 to eight in 2020 (Figure 6). Some fishermen have contributed with samples for many years. The annual number of shrimp samples from fishermen have similarly varied, as has the number of samples from the Coast Guard (Table 7). In 2015-2020, the number of samples has varied between 77 and 135. Sampling is fairly well distributed over the year (Tables 8, 9). Since 2006, between 50 and 90 % of the samples in Skagerrak have come from vessels < 15 m in length, while in the Norwegian Deep between 40 and 75 % of the samples have come from this vessel group (Figure 7). In some years, there are no samples from one or more quarters, this pertains specially to sampling from the larger vessels (> 15 m) in area 08 (Table 8). Within years, there have been some variation between areas (Tables 8, 9), where Skagerrak is the best sampled area (Figures 8-14). In some years, e.g. 2011, 2014 and 2017, coastal areas have been much better covered by sampling than off-shore areas. Since 2010, between 0.001 and 0.003 % of the landings have been sampled (Table 9).

Catch sampling from sorted catches

CRF data exist from Skagerrak since 2016 and from both areas since 2019. Norwegian length data from sorted catches from 2017-2020 were provided to the benchmark, but only data from 2019-2020 were used as input data for the model.

There are few vessels in the KRF, and the quarterly number of samples from sorted catches (with three catch fractions) are few compared to the number of samples from the sampling program based on unsorted catches (Tables 10, 11). Quarterly samples are missing for some year-quarter-area combinations (Table 11). The CRF vessels fishing in the Norwegian Deep fished in the whole area, while the CRF vessels in Skagerrak fished in much more restricted and localized areas (Figure 15). The LFD based on the CRF data (sorted catches) and the samples based on the unsorted catches are only partial in agreement (Figures 16, 17). The best agreement is seen for Skagerrak.

Discussion

Sampling of catches should reflect the spatial and temporal distribution of the fishery, as well as the fleet structure. This is especially important if vessels of different sizes fish differently, e.g. regarding selectivity of gears, and/or the stock structure varies spatially and temporally, e.g. regarding distribution of the different age groups. For the Norwegian shrimp fishery in Skagerrak and the Norwegian Deep, the fleet structure differs substantially between Skagerrak and the Norwegian Deep and the gear use (single, twin trawl) differs between coastal and off-shore areas. The use of selective devices deployed by the fleet is unfortunately unknown, as this information is not provided in the logbooks. The shrimp stock structure differs between Skagerrak and the Norwegian Deep with more juveniles in the former area in the first half of the year (Figures 16, 17) (Søvik and Thangstad 2021).

The first sampling program (sampling from unsorted catches) has had a better coverage of Skagerrak than the Norwegian Deep, but the fishery is also larger in the former area. The coverage of landings from the various vessel categories is skewed towards the smaller vessels; landings from the larger vessels often operating off-shore are less sampled than landings from the small vessels fishing on the stock along the coast. This disparity is largest in the Norwegian Deep where around 80 % of landings are taken by vessels ≥ 15 m, while never more than 60 % (most often less) of the samples come from this fleet segment.

The assessment model has a quarterly time step, necessitating samples from all four quarters. The temporal distribution of the sampling is good, but vessels ≥ 15 m are again underrepresented, and in some quarters in some years not covered at all.

With new requirements for data delivery to the assessment model, i.e. length samples from landings and discards separately, the Norwegian sampling program needs to be changed and hereafter based on data from the CRF. Presently too few samples are delivered from the CRF vessels, and there is a need for reinforcing this sampling program. This will be in place earliest in 2023 and requires information about the weight of the three different catch fractions, including discards. Discarding is illegal, thus trust between fishers and scientists is a prerequisite for the delivery of reliable data.

References

Hatlebrekke, H.H., Gundersen, S., Nedreaas, K., Vølstad, J.H. and Kolding, J. 2021. The Coastal Reference Fleet 2007-2019. Fleet composition, fishing effort and contributions to science. Rapport fra Havforskningen 2021-52. ISSN:1893-4536. 61 pp.

ICES 2016. Report of the Benchmark Workshop on Pandalus borealis in Skagerrak and Norwegian Deep Sea (WKPAND) 20–22 January 2016, Bergen, Norway. ICES CM 2016/ACOM:39. 71 pp.

ICES 2021. Joint NAFO/ICES *Pandalus* Assessment Working Group (NIPAG). ICES Scientific Reports. 3:22. 25 pp. https://doi.org/10.17895/ices.pub.7917

Søvik, G. and Thangstad, T.H. 2021. Results of the Norwegian Bottom Trawl Survey for Northern Shrimp (*Pandalus borealis*) in Skagerrak and the Norwegian Deep (ICES

Divisions 3.a and 4.a east) in 2021. NAFO SCR Doc. 21/001, Serial No. N7157. 38 pp. https://www.nafo.int/Portals/0/PDFs/sc/2021/scr21-001.pdf

Figures and tables

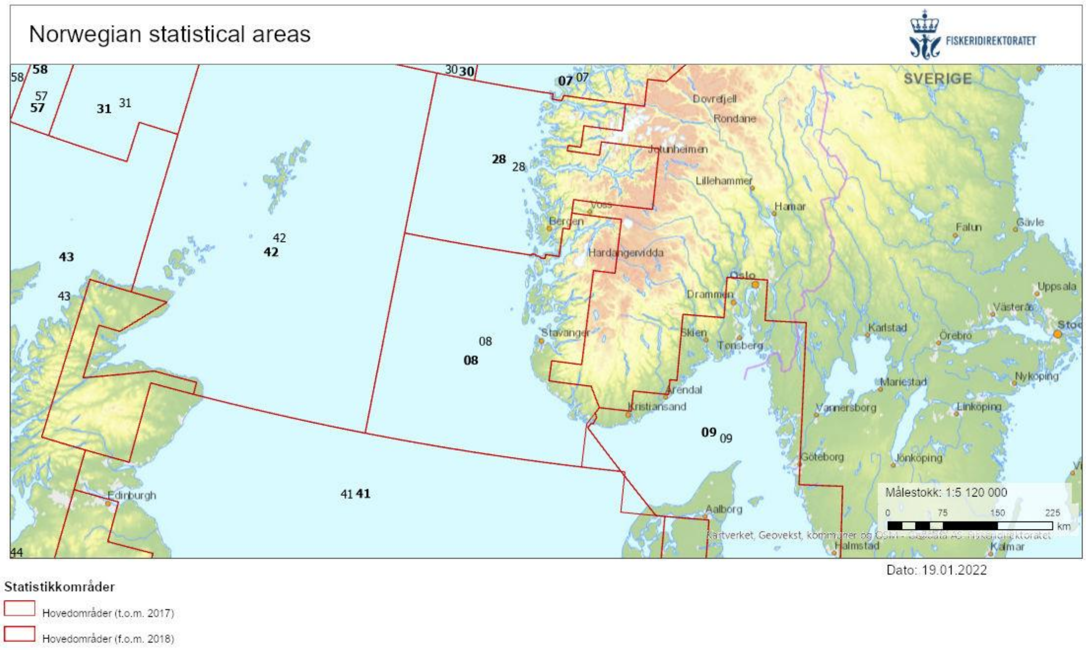


Figure 1a. Norwegian statistical areas ("Hovedområder") until 2017, and since 2018. Map from the Norwegian Directorate of Fisheries ([Fiskeri (fiskeridir.no)]).

Figure 1b. ICES Divisions ("ICES-områder") until 2017, and since 2018. Map from the Norwegian Directorate of Fisheries ([Fiskeri (fiskeridir.no)]).

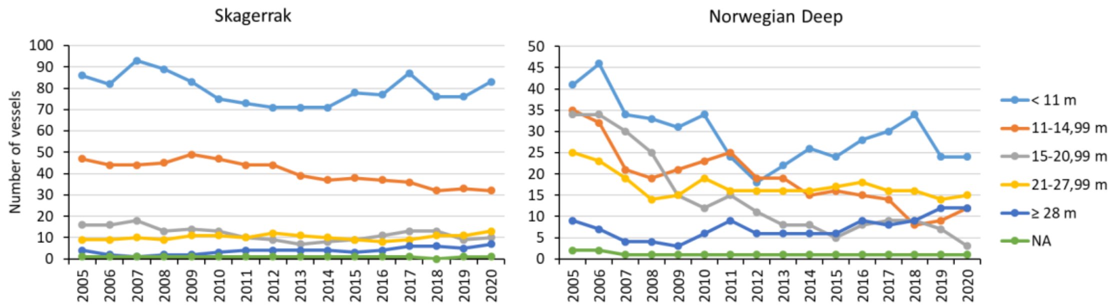


Figure 2. The Norwegian fleet of shrimp trawlers in Skagerrak and the Norwegian Deep: Number of vessels per length category and area, 2005-2020. Data from the Norwegian Directorate of Fisheries.

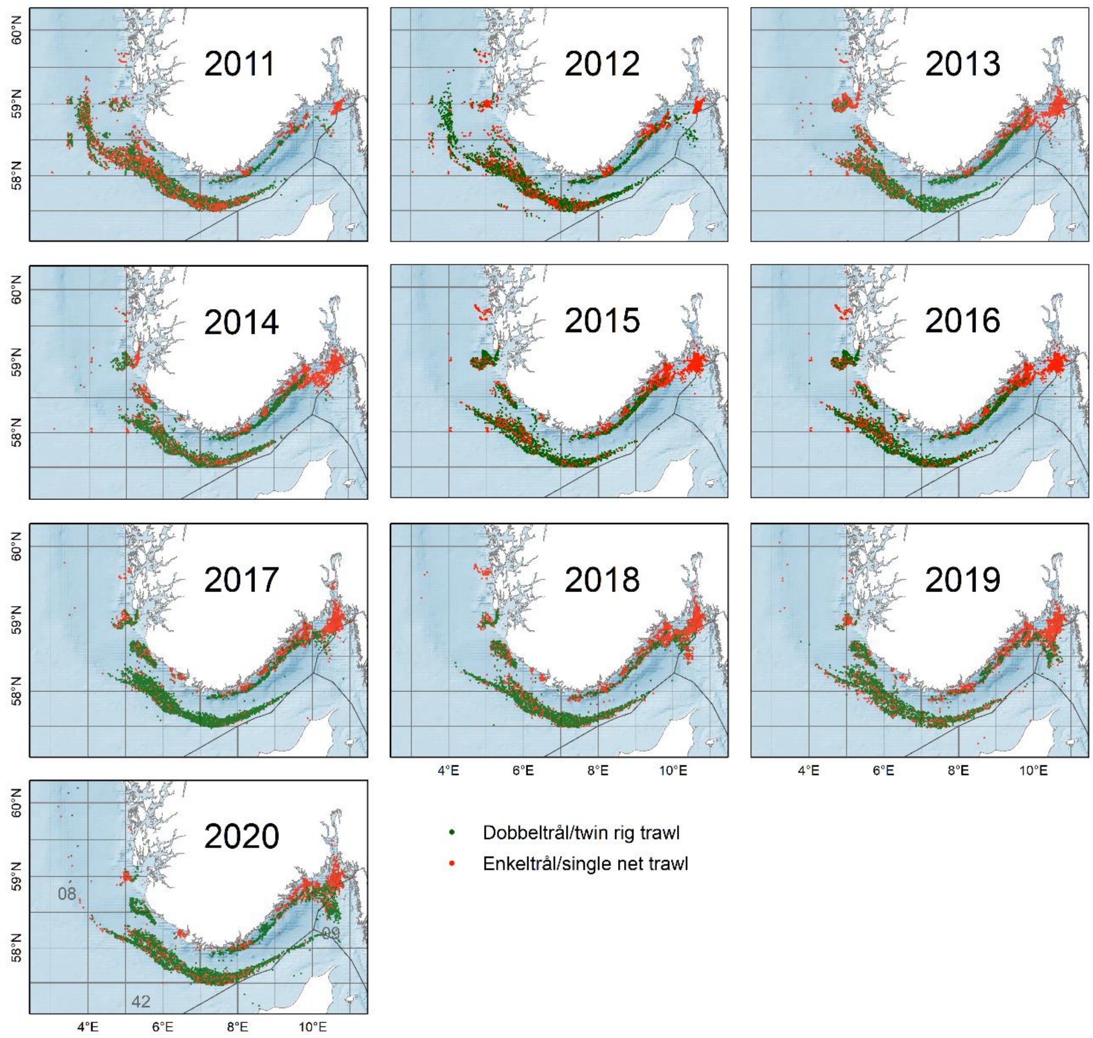


Figure 3. The distribution of the Norwegian shrimp fleet in Skagerrak and the Norwegian Deep, 2011-2020, based on registrations in electronic logbooks (available from 2011). In Skagerrak outside 4 nm (statistical area 09) and in the Norwegian Deep (statistical areas 08+28), all vessels respectively ≥ 12 m and ≥ 15 m are obliged to fill in logbooks. Data from the Norwegian Directorate of Fisheries.

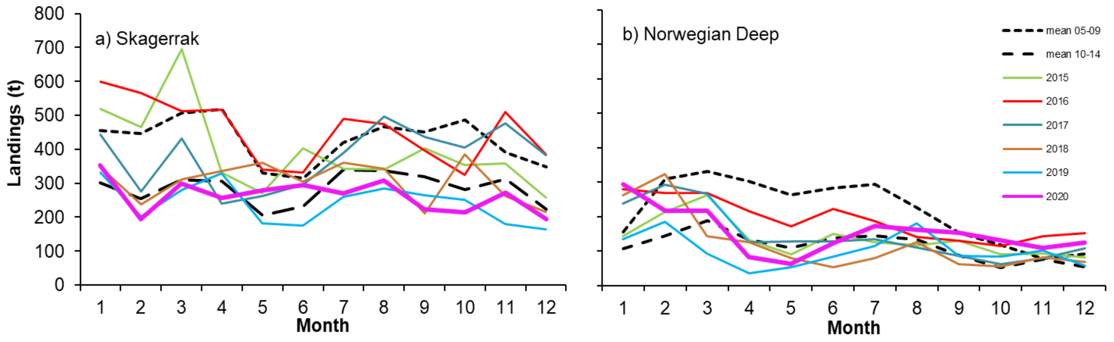


Figure 4. Monthly landings (tons) of the Norwegian shrimp fleet in a) Skagerrak and b) the Norwegian Deep in 2015-2020, and means 2005-2009 and 2010-2014. Data from the Norwegian Directorate of Fisheries.

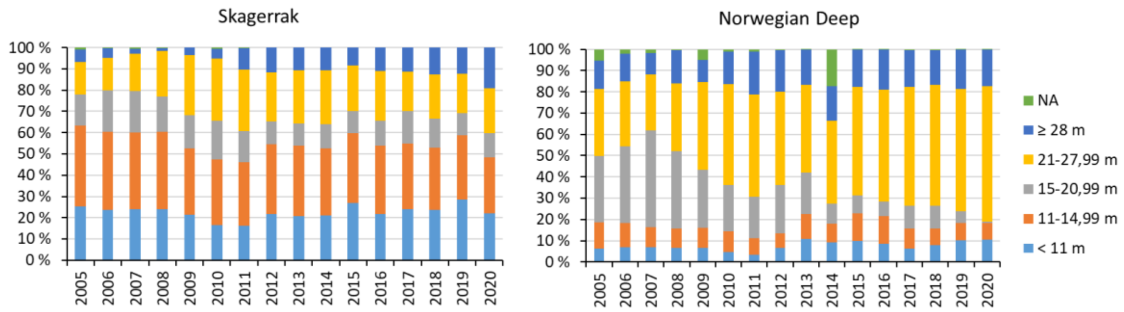


Figure 5. Landings per vessel length group (%) from the Norwegian shrimp fleet in Skagerrak (left) and the Norwegian Deep (right) in 2005-2020. Data from the Norwegian Directorate of Fisheries.

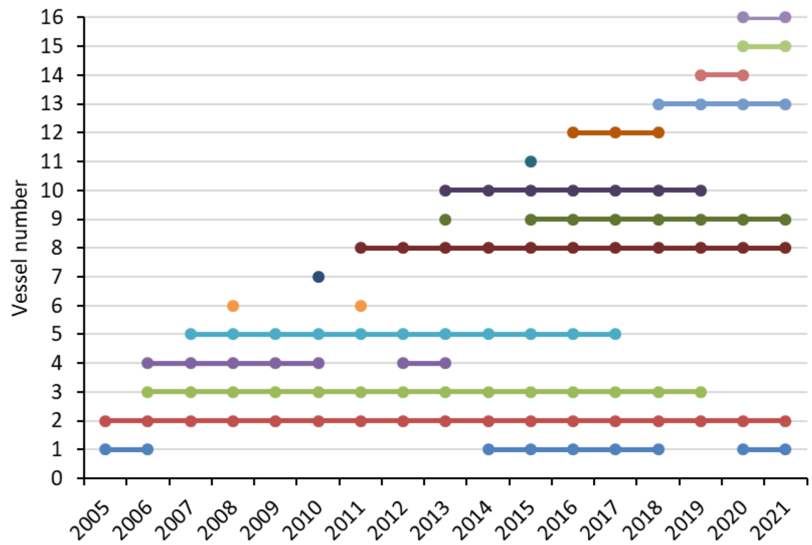


Figure 6. Number of fishermen (vessels) self-sampling their catches, per year, 2005-2021. Each line represents one vessel.

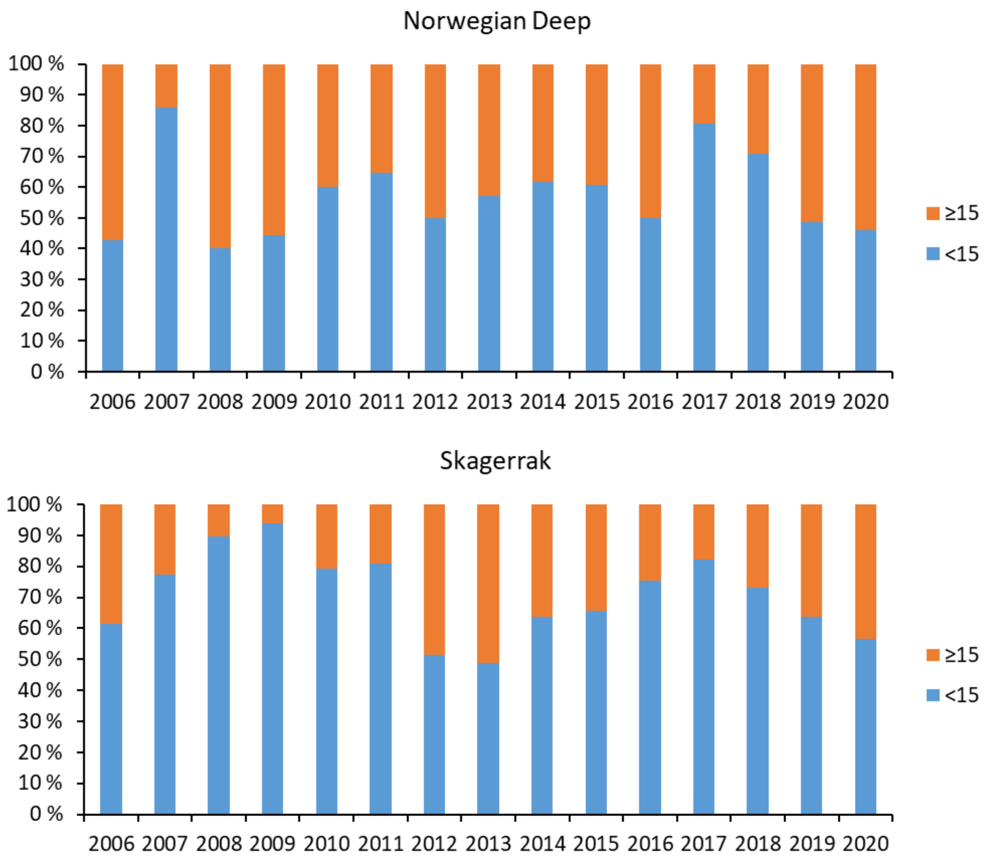


Figure 7. Annual percentages of shrimp samples from unsorted commercial catches from vessels < and ≥ 15 m length, 2006-2020, from the Norwegian Deep (upper) and Skagerrak (lower). Annual number of samples per area in Table 8.

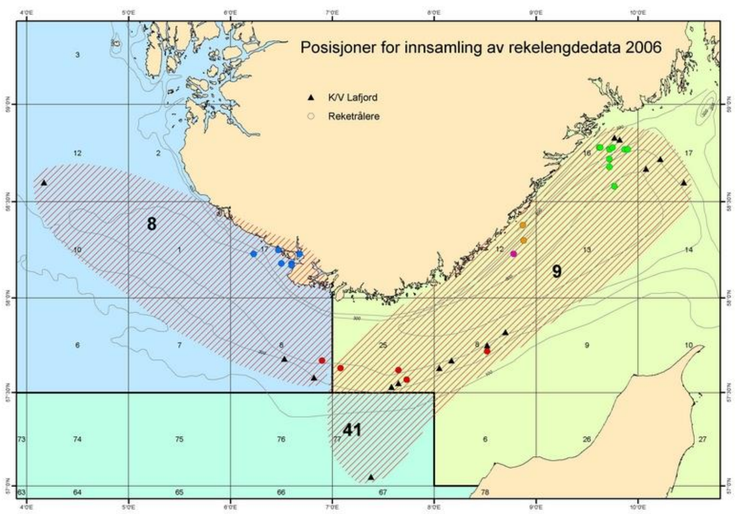


Figure 8. Positions of samples from unsorted commercial shrimp catches in 2006, where triangles indicate sampling by the Coast Guard K/V *Lafjord*, and dots indicate self-sampling by fishermen, the colours indicate different shrimp trawlers. Colour shading show the Norwegian statistical areas 08, 09 and 41.

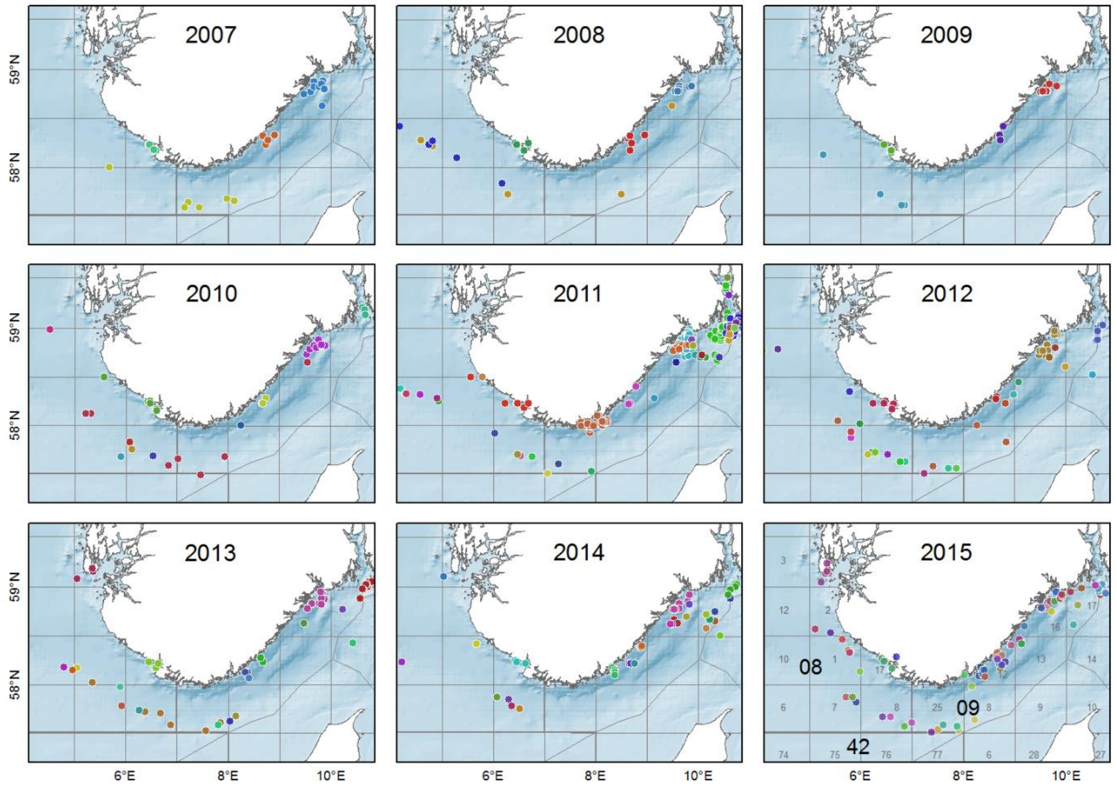


Figure 9. Positions of shrimp samples from unsorted commercial shrimp catches in 2007-2015, both self-sampling by fishermen and samples taken as part of Coast Guard inspections; colours indicate different vessels. Darker lines show the Norwegian statistical areas 08, 09 and 41.

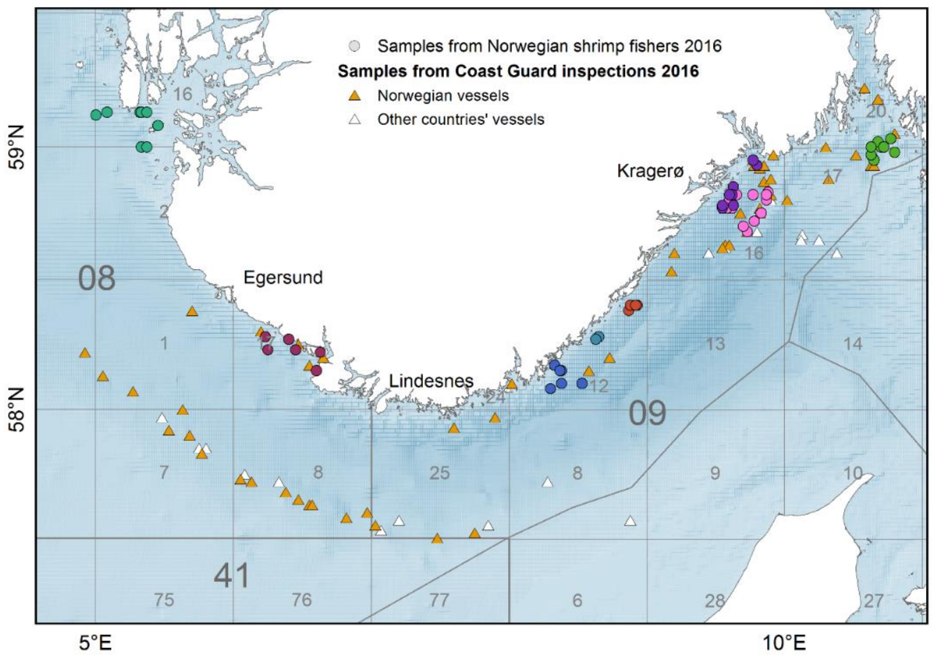


Figure 10. Positions of samples from unsorted commercial shrimp catches in 2016, where triangles indicate sampling by Coast Guard vessels and dots indicate self-sampling by fishermen. The colours indicate different shrimp trawlers. Darker lines show the Norwegian statistical areas 08, 09 and 41.

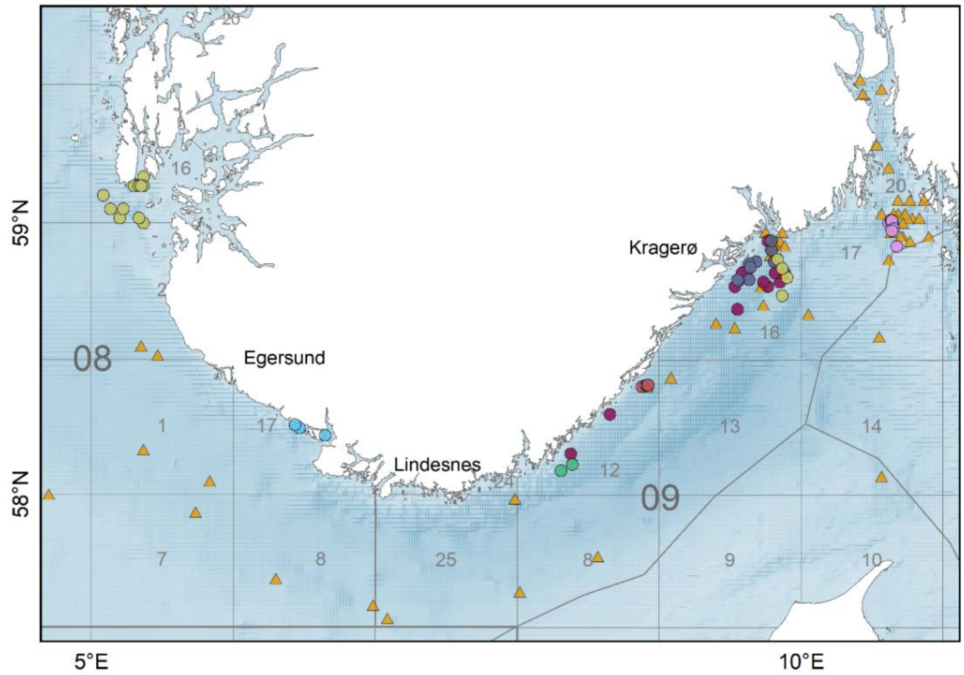


Figure 11. Positions of samples from unsorted commercial shrimp catches in 2017, where triangles indicate sampling by Coast Guard vessels and dots indicate self-sampling by fishermen. The colours indicate different shrimp trawlers. Darker lines show the Norwegian statistical areas 08 and 09.

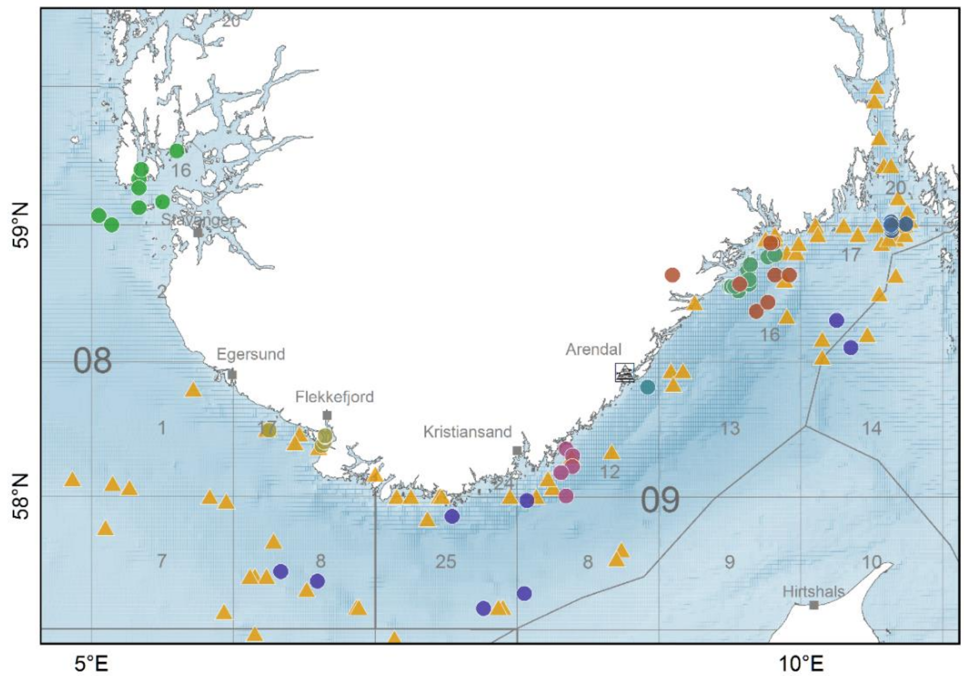


Figure 12. Positions of samples from unsorted commercial shrimp catches in 2018, where triangles indicate sampling by Coast Guard vessels and dots indicate self-sampling by fishermen. The colours indicate different shrimp trawlers. Darker lines show the Norwegian statistical areas 08 and 09.

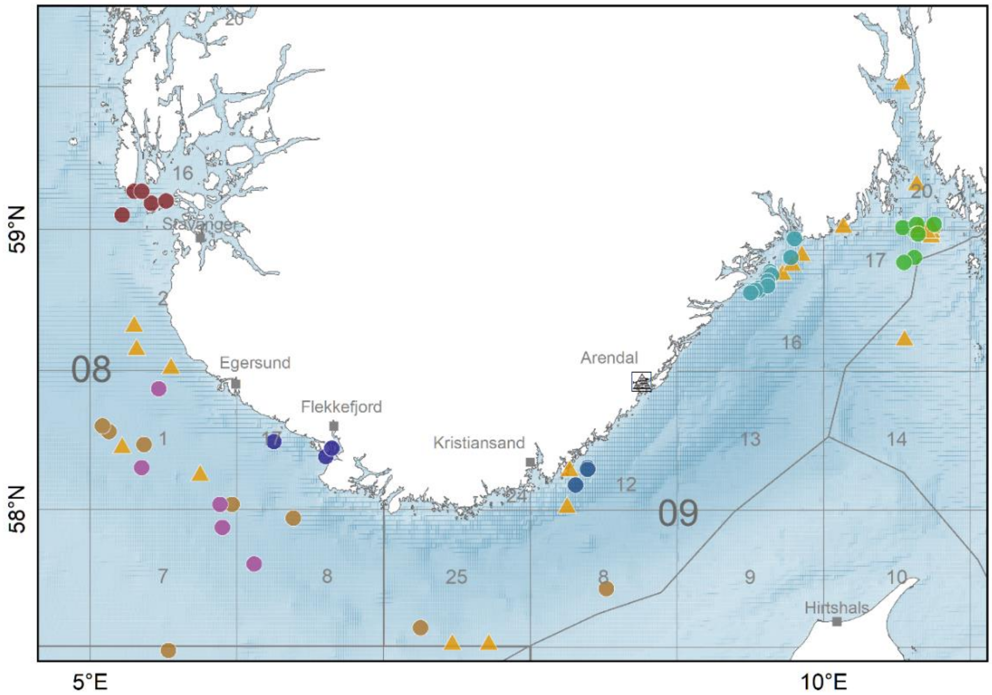


Figure 13. Positions of samples from unsorted commercial shrimp catches in 2019, where triangles indicate sampling by Coast Guard vessels and dots indicate self-sampling by fishermen. The colours indicate different shrimp trawlers. Darker lines show the Norwegian statistical areas 08 and 09.

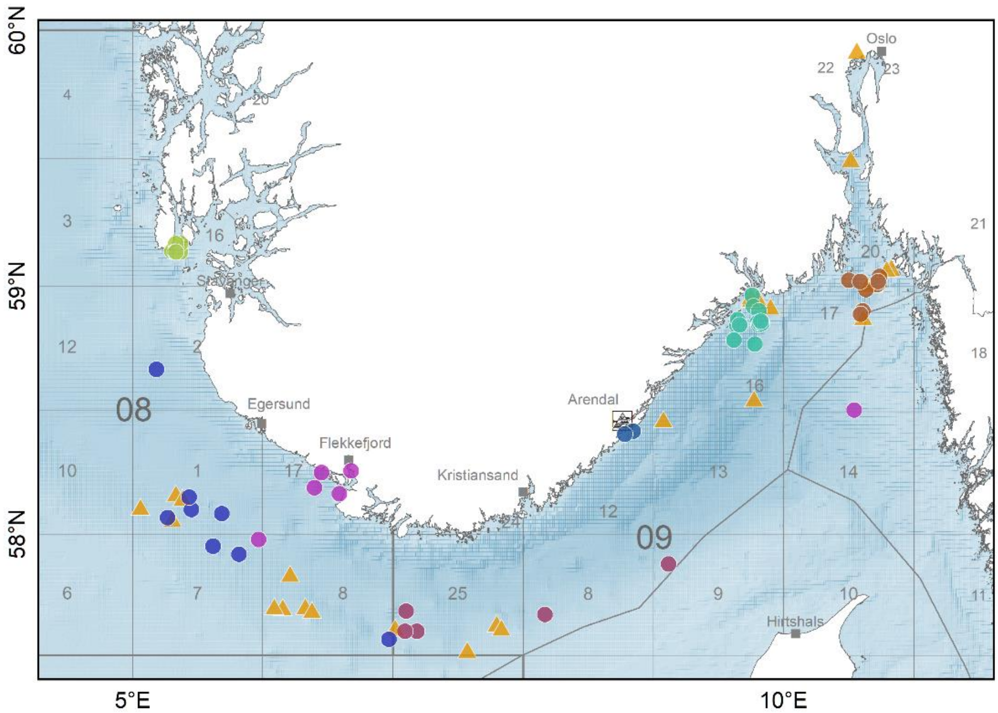


Figure 14. Positions of samples from unsorted commercial shrimp catches in 2020, where triangles indicate sampling by Coast Guard vessels and dots indicate self-sampling by fishermen. The colours indicate different shrimp trawlers. Darker lines show the Norwegian statistical areas 08 and 09.

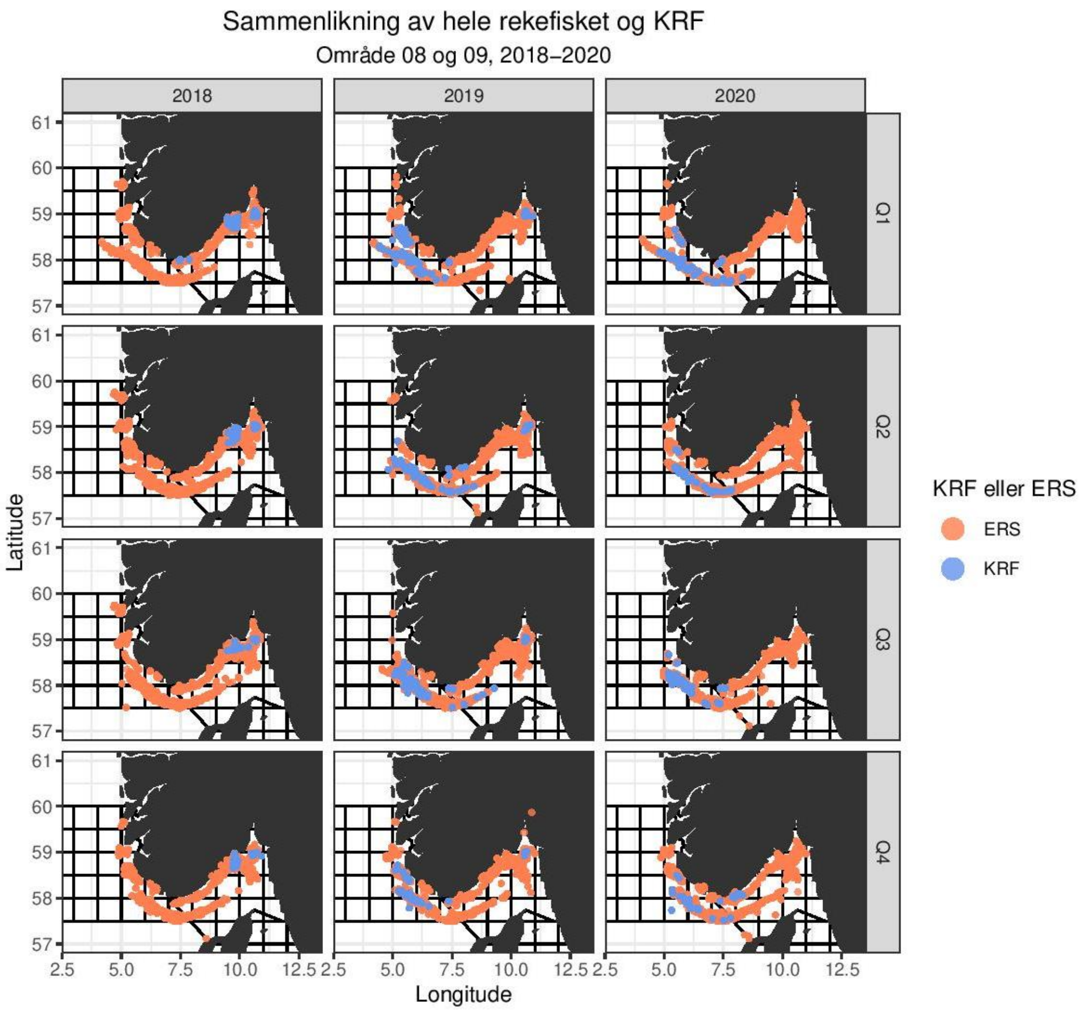


Figure 15. Positions of all reported trawl hauls from the shrimp trawlers in the Norwegian Coastal Reference fleet (blue dots) and positions of all trawl hauls in the electronic logbooks (orange dots), in 2018-2020, by quarter. All vessels in Skagerrak (outside 4 nm) and the Norwegian Deep with lengths of respectively ≥ 12 m and ≥ 15 m are obliged to fill in logbooks. ERS-data from the Norwegian Directorate of Fisheries. Figure by Emilie Rathe Knutsen.

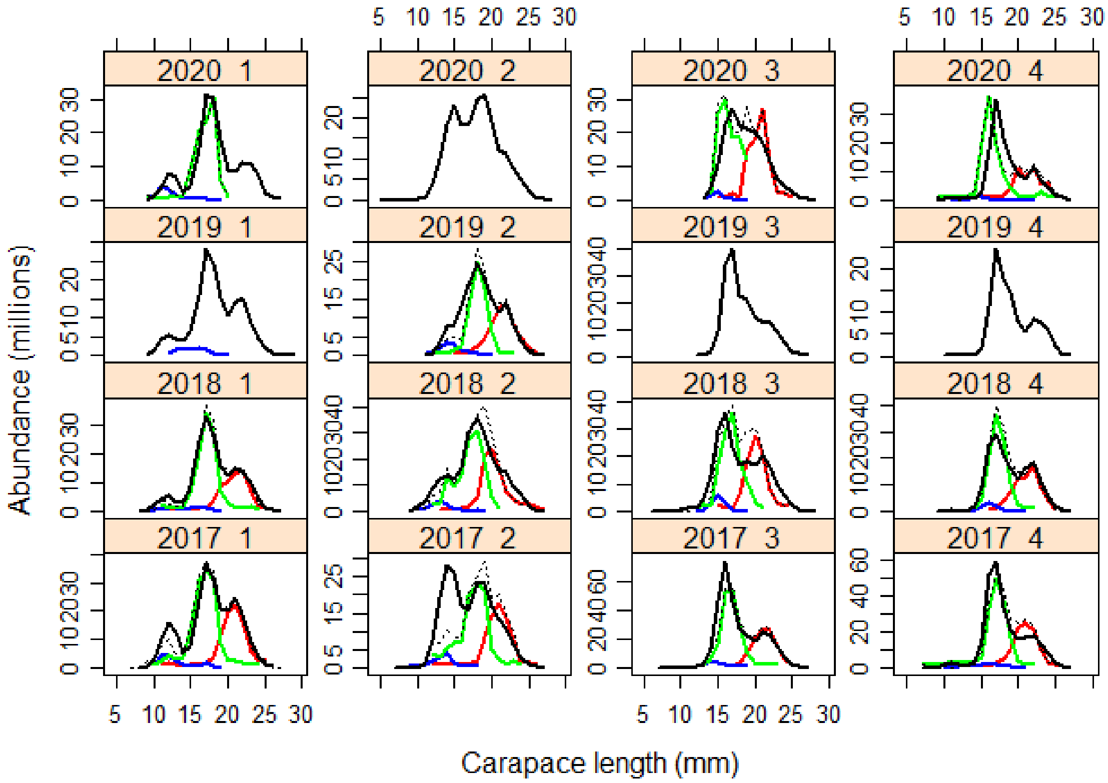


Figure 16. Quarterly length frequency distributions in Norwegian shrimp catches from Skagerrak in 2017-2020, based on sampling of unsorted catches (thick, black line) and sorted catches, where the blue line is discards, green is raw shrimp, red is large shrimp, and the thin, dotted black line is the sum of the three catch fractions. The modes represent age groups, where the first mode in quarter 1 is the 1-group, and the first mode in quarter 4 is the 0-group.

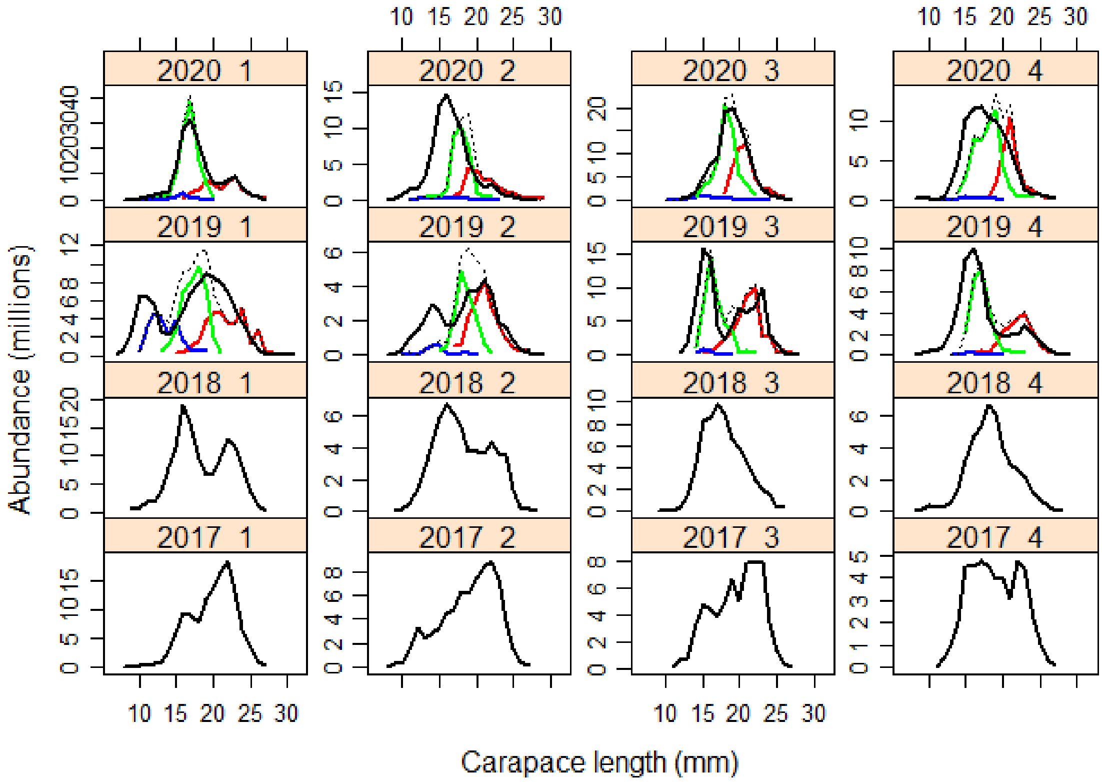


Figure 16. Quarterly length frequency distributions in Norwegian shrimp catches from the Norwegian Deep in 2017-2020, based on sampling of unsorted catches (thick, black line) and sorted catches, where the blue line is discards, green is raw shrimp, red is large shrimp, and the thin, dotted black line is the sum of the three catch fractions.

Table 1. Danish quarterly shrimp landings (corrected for boiling) and discards (in tons) from Skagerrak, 2009-2016, used for estimating Danish discard rates, which were applied to the Norwegian quarterly shrimp landings to estimate Norwegian discards for Skagerrak, in 2009-2016.

	quarter	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Landings	1	678	454	356	310	395	273	599	608
	2	630	334	443	317	598	562	642	403
	3	494	335	483	440	631	927	863	567
	4	633	235	318	387	402	670	604	418
Discards	1	12	40	17	18	81	128	72	30
	2	12	12	53	12	41	117	44	5
	3	11	5	32	18	34	57	54	0
	4	6	2	26	44	29	225	33	0

Table 2. The Norwegian fleet of shrimp trawlers in Skagerrak and the Norwegian Deep: number of vessels per length category and area, in 2005-2020. Data from the Norwegian Directorate of Fisheries.

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
	Skagerrak															
< 11 m	86	82	93	89	83	75	73	71	71	71	78	77	87	76	76	83
11-14,99 m	47	44	44	45	49	47	44	44	39	37	38	37	36	32	33	32
15-20,99 m	16	16	18	13	14	13	10	9	7	8	9	11	13	13	9	10
21-27,99 m	9	9	10	9	11	11	10	12	11	10	9	8	9	11	11	13
> 28 m	4	2	1	2	2	3	4	4	4	4	3	4	6	6	5	7
NA	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1
Total	163	154	167	159	160	150	142	141	133	131	138	138	152	138	135	146
	Norwegian Deep															
< 11 m	41	46	34	33	31	34	24	18	22	26	24	28	30	34	24	24
11-14,99 m	35	32	21	19	21	23	25	19	19	15	16	15	14	8	9	12
15-20,99 m	34	34	30	25	15	12	15	11	8	8	5	8	9	9	7	3
21-27,99 m	25	23	19	14	15	19	16	16	16	16	17	18	16	16	14	15
> 28 m	9	7	4	4	3	6	9	6	6	6	6	9	8	9	12	12
NA	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Total	146	144	109	96	86	95	90	71	72	72	69	79	78	77	67	67

Table 3. Norwegian quarterly and total landings (tons) per area, in 1988-2000. All landings since 2000 are corrected for boiling of the large shrimp. Data from the Norwegian Directorate of Fisheries.

	Skagerrak					**Norwegian Deep**				
	1	**2**	**3**	**4**	**Total**	**1**	**2**	**3**	**4**	**Total**
1988	989	827	761	470	3047	1686	1428	862	636	4612
1989	781	582	1144	649	3156	928	904	1077	510	3418
1990	621	681	864	841	3006	838	1098	706	503	3146
1991	837	630	1107	867	3441	926	682	626	430	2663
1992	1524	910	1040	782	4257	1139	816	648	343	2945
1993	877	625	1372	1214	4089	1038	698	1090	624	3449
1994	1107	1329	1037	915	4388	624	1060	506	237	2426
1995	1810	1156	1106	1110	5181	733	1056	603	447	2838
1996	1244	1049	1151	1714	5157	943	1043	400	367	2753
1997	1602	1391	1490	979	5461	988	1075	658	386	3107
1998	2191	1267	1814	1244	6515	1246	955	648	340	3189
1999	914	737	1218	1115	3985	905	647	727	472	2752
2000	1000	653	1203	877	3733	885	763	673	387	2709
2001	848	847	795	624	3114	1345	1265	1224	322	4156
2002	728	779	1443	930	3881	767	1251	1093	711	3822
2003	1084	765	1118	961	3927	1384	1227	984	663	4259
2004	1304	945	1455	1162	4867	1543	1336	1134	669	4681
2005	1205	976	1413	1018	4611	1316	1511	1040	481	4348
2006	1201	1318	1436	1440	5396	1055	966	934	318	3273
2007	1785	1184	1664	1564	6197	854	791	514	330	2489
2008	1696	1486	1500	1376	6058	504	780	690	230	2204
2009	1468	1105	944	1033	4550	592	560	474	186	1812
2010	775	709	709	624	2817	621	512	415	309	1856
2011	695	725	822	647	2890	582	625	580	124	1910

2012	1002	536	1159	1072	3768	364	278	264	178	1084
2013	1057	1001	1042	869	3969	365	305	381	159	1210
2014	1083	986	1568	1141	4779	455	339	329	221	1345
2015	1786	1071	1157	1032	5045	669	404	404	287	1763
2016	1788	1284	1452	1299	5823	884	661	495	442	2482
2017	1232	854	1427	1349	4861	864	423	359	270	1917
2018	956	1065	973	919	3913	786	280	287	226	1579
2019	873	734	866	634	3107	444	187	411	265	1307
2020	908	891	858	728	3386	773	285	517	388	1963

Table 4. Landings per vessel length group (tons) from the Norwegian shrimp fleet in Skagerrak and the Norwegian Deep, in 2005-2020. Data from the Norwegian Directorate of Fisheries

	2005	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
	Skagerrak															
< 11 m	1113	1227	1423	1379	914	430	433	769	781	942	1275	1189	1083	870	829	703
11-14,99 m	1685	1906	2135	2084	1327	808	813	1170	1239	1431	1548	1748	1401	1077	878	835
15-20,99 m	648	999	1163	970	677	469	386	387	388	511	491	644	694	498	299	362
21-27,99 m	672	792	1031	1217	1204	763	783	818	936	1130	1005	1267	849	759	537	675
> 28 m	259	240	143	85	144	126	276	418	393	485	397	602	509	465	356	607
NA	41	13	34	11	4	12	4	2	2	0	0	0	1	0	1	0
Total	4418	5177	5928	5746	4269	2608	2695	3564	3739	4500	4716	5450	4537	3670	2899	3182
	Norwegian Deep															
< 11 m	256	208	163	134	114	79	60	67	123	138	160	195	112	114	125	193
11-14,99 m	513	350	212	188	162	169	137	70	133	136	210	302	164	119	96	148
15-20,99 m	1302	1096	1057	741	475	377	353	230	223	141	139	157	188	156	68	12
21-27,99 m	1308	925	604	646	714	819	855	444	465	587	831	1207	992	834	699	1184
> 28 m	558	398	233	318	184	265	360	197	188	246	284	435	302	239	225	318
NA	219	63	40	11	84	19	21	5	3	264	1	1	9	6	2	4

Total	4156	3040	2309	2039	1733	1728	1787	1014	1134	1513	1625	2297	1766	1468	1214	1860

Table 5. Estimated quarterly discards (tons) from the Norwegian shrimp fleet in Skagerrak and the Norwegian Deep, in 2009-2020. Discards in 2009-2016 were estimated by applying Danish quarterly discard rates to Norwegian quarterly landings (corrected for weight loss due to boiling), while discards in 2017-2020 were estimated based on data from the Norwegian Coastal Reference fleet.

	Quarter	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Skagerrak	1	26	68	33	58	218	506	215	88	23	23	33	19
	2	21	25	87	20	69	205	74	17	24	24	34	11
	3	21	11	54	47	55	97	72	0	40	40	37	13
	4	10	5	53	122	62	383	57	0	27	27	19	7
Norw.Deep	1											44	16
	2											4	3
	3											5	8
	4											2	4

Table 6. Norwegian landings (in tons) of shrimp below minimum legal size (MLS) in Skagerrak and the Norwegian Deep, in 2016-2021. Data from the sales organization Fiskehav SA.

2016	2017	2018	2019	2020	2021
2	52	45	27	111	33

Table 7. Number of shrimp samples from unsorted catches from the Norwegian Coast Guard, commercial shrimp fishers and the Sea Surveillance Service (part of the Norwegian Fisheries Directorate) in 2005-2020.

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Coast Guard	3	14	1			4	17	14	17	15	42	62	49	51	22	28
Shrimp fishers	7	26	28	34	25	35	26	35	48	42	59	73	82	66	55	58
Sea Surveillance Service														11		4
Total	10	40	29	34	25	39	43	49	65	57	101	135	131	128	77	90

Table 8. Number of shrimp samples from unsorted catches by quarter, area and vessel length groups (< 15 m and ≥ 15 m), in 2006-2020.

Area	Quarter	Vessel size (m)	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Norwegian	1	<15		2	2	1	3	4	3	6	2	3	8	13	5	7	1
Deep		≥15	2			3			1	6	5	6	10	3	5	6	10
	2	<15	2	2	2	2	2	4	3	5	2	4	7	6	8	3	7
		≥15	3	1	4	2	3	1	3	2		2	9	2	1	5	2
	3	<15	2	1	1		2	1		3	2	6	2	2	4		2
		≥15	2		4			3	3	4		2		1	3	4	8
	4	<15	2	1	1	1	2	2	2	2	2	4	2	4	5	6	7
		≥15	1		1		3	2	1			1				2	
Skagerrak	1	<15	3	4	6	5	5	9	5	4	8	18	19	34	28	11	9
		≥15	1				1	1	4	5	4	5	4	4	6	5	5
	2	<15	5	5	4	4	7	5	4	4	6	8	20	13	12	7	14
		≥15	1	2			2	1	4	4	5	3	12	5	5	4	5
	3	<15	4	3	3	3	3	4	5	4	6	12	23	20	22	7	4
		≥15	2	2	1	1	1	1	5	6	4	9	5	4	10	4	10
	4	<15	4	5	4	3	4	3	3	6	8	10	11	15	9	3	3
		≥15	6	1	1		1	2	3	4	3	8	3	5	5	3	3
Total			40	29	34	25	39	43	49	65	57	101	135	131	128	77	90

Table 9. Number of shrimp samples, with number of specimens measured and sample weights from unsorted shrimp catches by area and quarter in 2005-2020. Quarterly landings (in tons) (uncorrected) and % of landings sampled are given.

Skagerrak	Norwegian Deep	Total

Year	Q	# samples	# shrimps	Sample weight (kg)	Landings (tons)	% of kg landed	# samples	# shrimps	Sample weight (kg)	Landings (tons)	% of kg landed	# samples	# shrimps	Sample weight (kg)	Landings (tons)	% of kg landed
2005	1	4	1249	6.0	1155	0.0005	1	306	1.9	1237	0.0001	5	1555	7.8	2392	0.0003
	2				935	0.0000	2	610	3.7	1420	0.0003	2	610	3.7	2355	0.0002
	3				1354	0.0000				978	0.0000				2332	0.0000
	4	3	1087	4.7	975	0.0005				452	0.0000	3	1087	4.7	1427	0.0003
	Total	7	2336	10.7	4419	0.0002	3	916	5.6	4087	0.0001	10	3252	16.3	8507	0.0002
2006	1	4	368	2.5	1152	0.0002	2	1235	6.5	979	0.0007	6	1603	9.0	2132	0.0004
	2	6	1715	11.0	1265	0.0009	5	1738	6.6	896	0.0007	11	3453	17.6	2161	0.0008
	3	6	1196	6.9	1378	0.0005	4	1210	5.0	867	0.0006	10	2406	11.8	2245	0.0005
	4	10	857	5.0	1382	0.0004	3	3209	13.9	295	0.0047	13	4066	19.0	1677	0.0011
	Total	26	4136	25.4	5177	0.0005	14	7392	32.1	3037	0.0011	40	11528	57.5	8214	0.0007
2007	1	4	1388	6.2	1708	0.0004	2	525	3.6	791	0.0005	6	1913	9.8	2499	0.0004
	2	7	1991	8.5	1132	0.0008	3	920	5.0	733	0.0007	10	2911	13.6	1865	0.0007
	3	5	1480	7.0	1592	0.0004	1	318	1.9	476	0.0004	6	1798	8.9	2068	0.0004
	4	6	1837	10.1	1496	0.0007	1	316	2.1	306	0.0007	7	2153	12.2	1802	0.0007
	Total	22	6696	31.8	5928	0.0005	7	2079	12.7	2307	0.0005	29	8775	44.5	8235	0.0005
2008	1	6	548	3.4	1608	0.0002	2	1691	9.4	466	0.0020	8	2239	12.8	2074	0.0006
	2	4	1507	8.1	1409	0.0006	6	1188	5.7	721	0.0008	10	2695	13.7	2130	0.0006
	3	4	1367	8.0	1423	0.0006	5	1186	5.7	639	0.0009	9	2553	13.7	2061	0.0007
	4	5	512	3.3	1305	0.0002	2	1510	8.5	213	0.0040	7	2022	11.8	1517	0.0008
	Total	19	3934	22.7	5744	0.0004	15	5575	29.2	2039	0.0014	34	9509	52.0	7783	0.0007
2009	1	5	1379	8.8	1377	0.0006	4	1014	6.3	545	0.0012	9	2393	15.1	1921	0.0008
	2	4	1107	5.9	1037	0.0006	4	911	5.5	519	0.0011	8	2018	11.4	1556	0.0007
	3	4	1154	6.4	886	0.0007				437	0.0000	4	1154	6.4	1323	0.0005
	4	3	764	5.0	969	0.0005	1	258	1.9	171	0.0011	4	1022	6.9	1140	0.0006
	Total	16	4404	26.2	4268	0.0006	9	2183	13.6	1672	0.0008	25	6587	39.8	5940	0.0007

2010	1	6	1272	8.5	714	0.0012	3	571	4.9	576	0.0008	9	1843	13.4	1291	0.0010
	2	9	2067	12.3	654	0.0019	5	961	6.5	471	0.0014	14	3028	18.8	1126	0.0017
	3	4	1307	7.9	654	0.0012	2	269	1.5	377	0.0004	6	1576	9.4	1032	0.0009
	4	5	1337	8.2	575	0.0014	5	1113	8.7	284	0.0031	10	2450	16.9	860	0.0020
	Total	24	5983	37.0	2598	0.0014	15	2914	21.6	1709	0.0013	39	8897	58.6	4308	0.0014
2011	1	10	3002	19.1	648	0.0030	4	946	7.4	540	0.0014	14	3948	26.5	1188	0.0022
	2	6	1788	7.5	676	0.0011	5	1001	7.5	580	0.0013	11	2789	15.0	1256	0.0012
	3	5	1515	7.4	766	0.0010	4	999	8.1	538	0.0015	9	2514	15.4	1305	0.0012
	4	5	1499	8.0	603	0.0013	4	1164	6.7	115	0.0058	9	2663	14.7	718	0.0020
	Total	26	7804	42.0	2693	0.0016	17	4110	29.6	1773	0.0017	43	11914	71.6	4466	0.0016
2012	1	9	2779	12.4	947	0.0013	4	891	5.4	343	0.0016	13	3670	17.8	1289	0.0014
	2	8	2459	8.1	507	0.0016	6	1498	8.5	258	0.0033	14	3957	16.7	765	0.0022
	3	10	2946	12.5	1097	0.0011	3	662	3.6	243	0.0015	13	3608	16.1	1340	0.0012
	4	6	1681	8.0	1014	0.0008	3	918	4.3	165	0.0026	9	2599	12.3	1178	0.0010
	Total	33	9865	41.0	3564	0.0012	16	3969	21.8	1009	0.0022	49	13834	62.8	4572	0.0014
2013	1	9	2627	12.3	996	0.0012	12	2906	15.2	342	0.0044	21	5533	27.5	1337	0.0021
	2	8	2295	9.7	943	0.0010	7	1992	8.8	285	0.0031	15	4287	18.5	1229	0.0015
	3	10	2654	12.3	981	0.0012	7	2125	9.2	356	0.0026	17	4779	21.5	1338	0.0016
	4	10	2557	13.9	818	0.0017	2	598	3.4	149	0.0023	12	3155	17.3	968	0.0018
	Total	37	10133	48.2	3739	0.0013	28	7621	36.6	1132	0.0032	65	17754	84.8	4871	0.0017
2014	1	12	3382	15.4	1020	0.0015	7	1713	7.7	423	0.0018	19	5095	23.1	1443	0.0016
	2	11	3010	10.9	930	0.0012	2	598	2.4	314	0.0008	13	3608	13.3	1244	0.0011
	3	10	3019	12.1	1476	0.0008	2	616	2.3	306	0.0008	12	3635	14.4	1782	0.0008
	4	11	3164	14.0	1074	0.0013	2	614	2.8	206	0.0013	13	3778	16.7	1280	0.0013
	Total	44	12575	52.3	4500	0.0012	13	3541	15.2	1249	0.0012	57	16116	67.5	5749	0.0012
2015	1	23	6482	29.3	1678	0.0017	9	1993	12.1	618	0.0020	32	8475	41.4	2296	0.0018
	2	11	3291	11.7	1006	0.0012	6	1544	7.6	373	0.0021	17	4835	19.4	1379	0.0014
	3	21	6054	26.6	1087	0.0024	8	2287	11.7	373	0.0031	29	8341	38.3	1460	0.0026

	4	18	5583	27.8	969	0.0029	5	1464	7.3	265	0.0028	23	7047	35.1	1234	0.0028
	Total	73	21410	95.4	4741	0.0020	28	7288	38.8	1628	0.0024	101	28698	134.2	6369	0.0021
2016	1	23	7188	34.5	1677	0.0021	18	4954	31.8	817	0.0039	41	12142	66.3	2494	0.0027
	2	32	9082	40.0	1189	0.0034	16	3922	20.5	612	0.0033	48	13004	60.5	1801	0.0034
	3	28	7424	34.5	1362	0.0025	2	406	2.4	459	0.0005	30	7830	36.8	1821	0.0020
	4	14	3711	20.5	1221	0.0017	2	602	4.0	410	0.0010	16	4313	24.5	1631	0.0015
	Total	97	27405	129.6	5449	0.0024	38	9884	58.6	2297	0.0026	135	37289	188.2	7746	0.0024
2017	1	38	10111	51.9	1151	0.0045	16	3383	23.8	795	0.0030	54	13494	75.7	1945	0.0039
	2	18	4601	19.0	798	0.0024	8	1400	8.0	382	0.0021	26	6001	27.0	1180	0.0023
	3	24	5960	27.2	1325	0.0021	3	494	2.8	331	0.0009	27	6454	30.0	1656	0.0018
	4	20	5211	27.3	1264	0.0022	4	729	4.3	250	0.0017	24	5940	31.6	1515	0.0021
	Total	100	25883	125.4	4537	0.0028	31	6006	38.9	1758	0.0022	131	31889	164.2	6296	0.0026
2018	1	34	8008	42.8	897	0.0048	10	2344	13.8	729	0.0019	44	10352	56.6	1625	0.0035
	2	17	4323	19.8	1000	0.0020	9	2253	11.0	260	0.0042	26	6576	30.8	1260	0.0024
	3	32	7946	36.8	913	0.0040	7	1571	7.6	266	0.0028	39	9517	44.3	1179	0.0038
	4	14	3476	19.0	861	0.0022	5	1515	7.7	207	0.0037	19	4991	26.7	1068	0.0025
	Total	97	23753	118.3	3670	0.0032	31	7683	40.1	1462	0.0027	128	31436	158.4	5132	0.0031
2019	1	16	4426	24.0	814	0.0030	13	3027	15.8	412	0.0038	29	7453	39.8	1226	0.0032
	2	11	3039	15.4	685	0.0023	8	1989	10.2	173	0.0059	19	5028	25.6	858	0.0030
	3	11	2895	14.3	808	0.0018	4	1123	5.4	381	0.0014	15	4018	19.6	1189	0.0016
	4	6	1728	10.4	593	0.0018	8	2169	10.6	246	0.0043	14	3897	21.0	839	0.0025
	Total	44	12088	64.1	2899	0.0022	33	8308	41.9	1212	0.0035	77	20396	106.0	4111	0.0026
2020	1	14	3851	20.9	845	0.0025	11	3389	17.6	728	0.0024	25	7240	38.5	1573	0.0024
	2	19	5666	25.5	823	0.0031	9	2998	11.0	268	0.0041	28	8664	36.5	1092	0.0033
	3	14	4089	20.8	797	0.0026	10	2689	13.7	488	0.0028	24	6778	34.4	1286	0.0027
	4	6	1803	9.9	700	0.0014	7	2565	11.9	371	0.0032	13	4368	21.8	1071	0.0020
	Total	53	15409	77.1	3165	0.0024	37	11641	54.2	1856	0.0029	90	27050	131.3	5021	0.0026

Table 10. Number of shrimp vessels in the Norwegian Coastal Reference fleet delivering catch data and shrimp samples from sorted catches, per area and year, 2015-2020.

Area	**2016**	**2017**	**2018**	**2019**	**2020**
Skagerrak	3	2	2	2	2
Norwegian Deep				2	2

Table 11. Number of shrimp samples from the Norwegian Coastal Reference fleet, by area (Skagerrak 3.a, and the Norwegian deep, 4.a), quarter and catch fraction (boiled, raw, and juveniles which are either landed or discarded), in 2017-2020.

Area		**4.a**				**3.a**			
Quarter		**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
2017	boiled					9	9	6	6
	raw					8	9	6	6
	discards					8	4	6	7
2018	boiled					8	5	5	5
	raw					8	5	5	5
	discards					8	5	5	5
2019	boiled	2	2	1	3		11		
	raw	2	2	1	3		9		
	discards	1	2	1	3	1	11		
2020	boiled	3	2	4	3	3		1	5
	raw	3	2	4	3	3		1	5
	discards	3	2	3	3	3		1	5

WD 3: Stock assessment of Northern shrimp (Pandalus borealis) in divisions 3.a and 4.a East (Skagerrak and Kattegat and northern North Sea in the Norwegian Deep)

By

Massimiliano Cardinale, Alessandro Orio, Mikaela Bergenius-Nord, Katja Nören and Francesco Masnadi

The majority of the contents of this WD are included Section 3.4.

WD 4 Natural Mortality for pra.27.3a4

By Massimiliano Cardinale and Mikaela Bergenius-Nord

The natural mortality rate (M) of fish populations is one of the most important parameters for population dynamics and stock assessment models. Unfortunately, it is also one of the most difficult parameters to estimate. For this benchmark assessment a pool of methodologies can be considered to assess the impact of M on the assessment.

The Barefoot Ecologist's Toolbox (http://barefootecologist.com.au/shiny_m) can be used to derive different values of single M or to derive composite M value weighting different methods. This toolbox, developed by Jason Cope, provides a straightforward method for obtaining the estimated value of natural mortality from a range of life-history based methods (different life-history input requirement).

In the Table 1 a summary of the input and output of all methods considered in the Toolbox divided by different input requirements (Input Categories). The VB parameter were taken from last assessment model (CIT) while the age at Mat was derived from L50% (1.975 cm: Mikaela document value) using standard VB equation.

Table 1: Natural mortality (M) from a range of life-history based methods.

	Methods	Reference	Input Categories	Input parms	Input Value	M output Value
Vector by age	Chen-Wat	*Chen & Watanabe 1989*	Amax & VBGP	Age, k, t0	Age:8 k:0.39; t0: 0	fig X1
Single M value*	Then_nls	*Then et al. 2015*	Amax	maximum age	8	0.72
	Then_lm	*Then et al. 2015*	Amax	maximum age	8	0.68
	Hamel_Amax	*Hamel. 2015; Hamel in pres.*	Amax	maximum age	8	0.67
	Hamel_k	*Hamel. 2015; Hamel in pres.*	VBGP	k	0.39	0.68
	Jensen_k1	*Jensen 1997*	VBGP	k	0.39	0.58
	Jensen_k2	*Jensen 1997*	VBGP	k	0.39	0.62
	Roff	*Roff 1984*	VBGP & Mat	k, age at maturity	k:0.39; 2.93	0.6
	Jensen_Amat	*Jensen 1996*	Mat	age at maturity	2.93	0.56
	Ri_Ef_Amat	*Rikhter & Efanov 1976*	Mat	age at maturity	2.93	0.54

**Single M values are translated in age-based vector using the proportion between ages in Chen-Watanabe vector by age.*

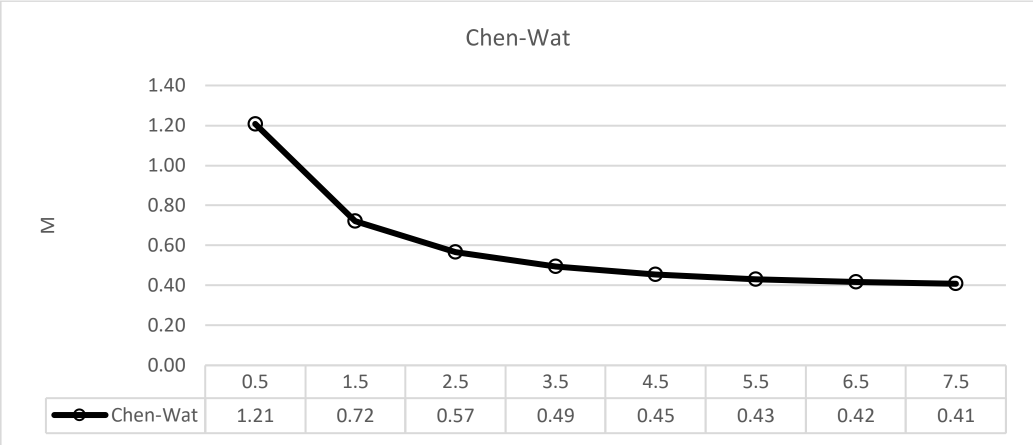
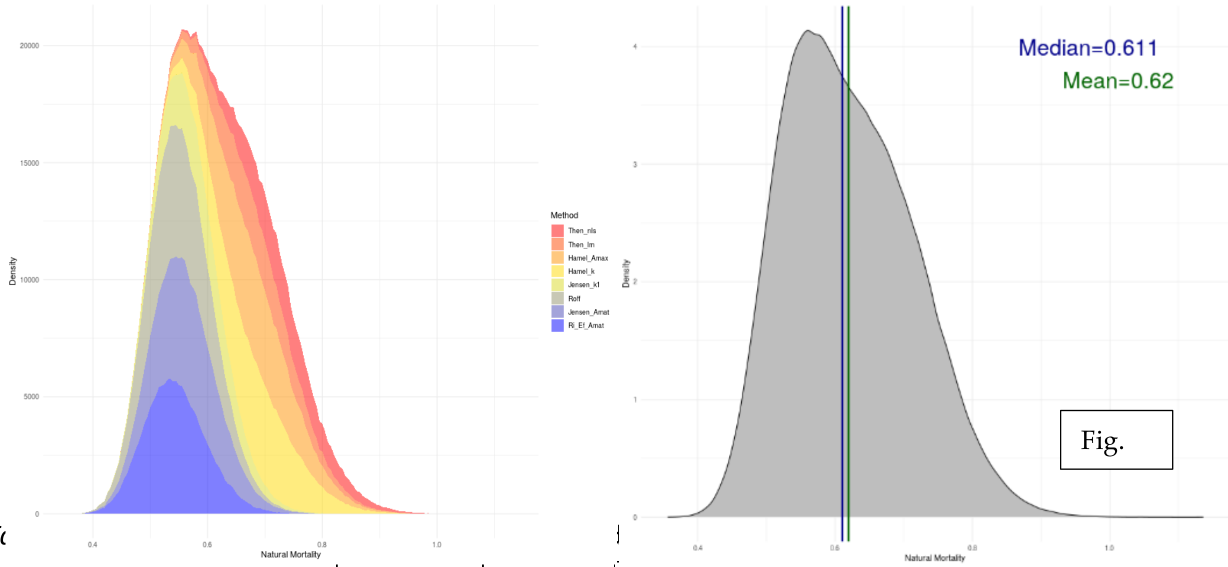


Figure 1

Final derive composite M (Fig. X2) was calculated using the 9 single Ms methods applying a CV of 0.1 to add additional uncertainty to the point estimates. Median, 5% and 95% CIs values were calculated (Tab 2).



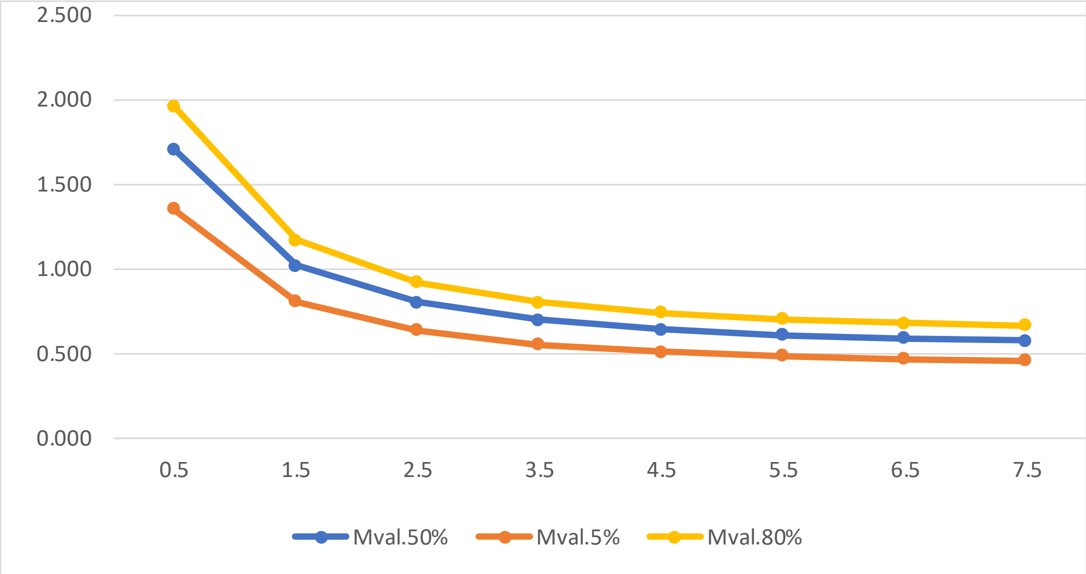
Tc

Mval.50%	CI 5%	CI 95%
0.61	0.48	0.78

For the ensemble grid, three values have been selected to represent uncertainty around natural mortality (lower, upper and median level) to be tested in the assessment. Those were CI5%, 50% and 80%. CI 80% was used instead of CI 95% to avoid the long tail of the M distribution. Single values were translated in age-based vector using the proportion between ages in Chen-Watanabe vector by age (fig X1) considering age 5.5 as reference age (proportion 100%). The three Ms (fig X3) will be treated as alternative hypothesis in the context of the ensemble approach (M as one dimension of the ensemble grid).

Age	Chen-Wat	Proportion to age 5.5 in CW method	Median	CI 5%	CI 80%
0.5	1.21	280%	1.709	1.356	1.964

1.5	0.72	167%	1.019	0.809	1.171
2.5	0.57	131%	0.800	0.635	0.920
3.5	0.49	114%	0.699	0.554	0.803
4.5	0.45	105%	0.643	0.511	0.739
5.5	**0.43**	**100%**	**0.611**	**0.485**	**0.702**
6.5	0.42	97%	0.590	0.468	0.679
7.5	0.41	95%	0.577	0.458	0.664



Nat Mort references

Then_nls, Then_lm

Then, A.Y., J.M. Honeig, N.G. Hall, D.A. Hewitt. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science 72(1): 82-92.

Hamel_Amax, Hamel_k

Hamel, O.S., 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES Journal of Marine Science 72, 62-69. https://doi.org/10.1093/icesjms/fsu131

Owen Hamel (in. prep; owen.hamel@noaa.gov)

Chen-Wat

Chen, S. and S. Watanabe. 1989. Age Dependence of Natural Mortality Coefficient in Fish Population Dynamics. Nippn Suisan Gakkaishi 55(2): 205-208.

Jensen_k1, Jensen_k2, Jensen_Amat

Jensen, A.L. 1996. Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. Can. J. Fish. Aquat. Sci. 53: 820-822.

Jensen, A.L. 1997. Origin of the relation between K and Linf and synthesis of relations among life history parameters. Can. J. Fish. Aquat. Sci. 54: 987-989.

Roff

Roff, D. A. 1984. The evolution of life history parameters in teleosts. Can. J. Fish. Aquat. Sci. 41: 989-1000.

Ri_Ef_Amat

Rikhter, V.A., Efanov, V.N., 1976. On one of the approaches to estimation of natural mortality of fish populations. ICNAF Res. Doc. 79/VI/8, 12.

WD 5 Sampling design and estimation of commercial catches of northern shrimp *Pandalus borealis* by the Swedish bottom otter trawl fisheries in the 2016- 2020

Sampling design and estimation of commercial catches of northern shrimp *Pandalus borealis* by the Swedish bottom otter trawl fisheries in the 2016-2020

Nuno Prista[a], Hongru Zhai[a], Katja Norén[a], Sofia Carlshamre[a], Annica de Groote[b]

[a]Department of Aquatic Resources, Swedish University of Agricultural Sciences, Sweden
[b]Department of Energy and Technology, Swedish University of Agricultural Sciences, Sweden

Abstract

The present working document describes work done in late 2021 to improve the 2016-2020 Swedish estimates of commercial catches of northern shrimp *Pandalus boreali*s supplied to NIPAG assessments. During the study period, the sampling of *Pandalus* fisheries carried out under the Swedish national data collection programme (EU fisheries data collection regulation) took place in relatively probabilistic way but commercial catch estimates have not, to date, properly accounted for the probabilistic nature of the sampling design. The present working document provides details on sampling design and onboard protocols, data storage and quality checks, and new estimation routines that more explicitly incorporate aspects of the design such as stratification, clustering, sampling methods and sampling probabilities of the data collected. The full array of new estimates was made available to WKPRAWN with a few examples of results obtained being illustrated in the present document. The new estimation routines open the possibility of evaluating biases and precision in data used in assessment, indicating steps forward that can still be taken to further improve shrimp estimates of commercial catches entering assessment.

1. Introduction

Prior to the present Benchmark Workshop on *Pandalus* stocks (WKPRAWN 2022) the main estimates of Swedish catches provided to NIPAG were live weights (in kg), and numbers@length of *Pandalus borealis* caught by the Swedish fleet in each year. In parallel, information was also provided on discard rate. From 2012 both types of estimates consisted of point estimates and resulted from unweighted aggregations of raw sampling data raised to haul level and then expanded to quarterly*area fleet level using simple ratios (e.g., landings in the fishery/landings in hauls sampled). In 2019-2020 discard rates were used to calculate discard from landings and numbers at length for unsorted samples were raised to catch. Such unweighted simple ratio estimators ignored the underlying probabilistic sampling design that originated the data and relied on some unchecked assumptions, increasing the risk of errors in estimates provided to assessment.

Over the last decade, the ICES commercial sampling community has been increasingly focused on the implementation of statistically sound survey designs in European surveys of commercial fisheries. That focus originated from work carried out by an extensive number of ICES expert groups (e.g., ICES 2005, 2011, 2012, ...) that highlighted that probability-based sampling was fundamental to, among other things, evaluations of bias and precision of commercial catch estimates annually used in stock assessments (ICES 2008, 2009). These efforts resulted in the introduction of the requirement for statistically sound sampling in EU fisheries data collection regulations making it a national obligation[1] and provided a final motivation for the implementation of probabilistic sampling designs in member states that were yet to do so, increasing the collection of representative data. Probabilistic collection of data does not, however, necessarily ensure that unbiased estimates are obtained, and precision is correctly quantified. For those to take place, the characteristics of the design, namely its stratification, multi-stage nature, its sampling methods and sampling probabilities need to be explicitly incorporated into the estimation process.

The present Benchmark Workshop on Pandalus stocks (WKPRAWN 2022) requested the disaggregation of the previous estimates (weight and numbers@length) by quarter, division and catch fraction (landings and discards). Such disaggregation is difficult to obtain from older data where documentation of sampling and estimation procedures used are frequently lacking (see Noren, Description of methodology behind compilation of Swedish data on Pandalus borealis (1908-2020) used in pandalus benchmark conducted in January 2022) but can be attempted for more recent years where data collection is better documented and data itself more quality checked. Such an attempt provides an opportunity to revisit the estimates previously provided to assessment and update them, re-calculating them with improved estimators that explicitly incorporate the probabilistic sampling design that underlies the data.

The present work details the main probabilistic elements of the sampling design used in the sampling of Swedish fisheries targeting *Pandalus borealis* over the period 2016-2020 and updates estimates provided to NIPAG by using design-based estimators. The 2016-2020 period was chosen based on aspects such as the higher level of implementation of probabilistic sampling, the higher degree of quality checking of survey data, and the existence of better data documentation, but the analysis described here can potentially be extended a few years more back in time in a future occasion.

2. Brief overview of the Swedish *Pandalus borealis* fishery during the period and onboard processing of catch by fishers

Overview of the Swedish Pandalus borealis fishery during 2016-2020

The Swedish fishery takes place in Kattegat (27.3.a.21), Skagerrak (27.3.a.20) and North Sea (27.4.a), with a clear dominance of Skagerrak (27.3.a.20). Two shrimp fisheries are considered to exist in the area: a shrimp fishery with bottom otter trawls equipped with fish tunnel (SWE_OTB_CRU_32-69_0_0) and a shrimp fishery with bottom otter trawls are not equipped with fish tunnel (SWE_OTB_CRU_32-69_2_22). Both fisheries are equipped with a similar shrimp grid but given its partial-fish orientation, over the study period fishery SWE_OTB_CRU_32-69_0_0 tended to take place more offshore and use otter twin trawls, having larger by-catches and discards of fish and other commercial species, the fishery SWE_OTB_CRU_32-69_2_22 (which tended to be more inshore, use otter single trawls, and have more limited by-catches of fish and other commercial species).

[1] Commission implementing decision (EU) 2016/1251 of 12 July 2016 adopting a multiannual Union programme for the collection, management and use of data in the fisheries and aquaculture sectors for the period 2017-2019

Table: Number of Swedish vessels and trips (in parentesis) targeting shrimp by fishery and ICES subdivision (source: 2016-2020 logbooks).

	SWE_OTB_CRU_32-69_0_0			**SWE_OTB_CRU_32-69_2_22**		
	27.3a.21	27.3.a.20	27.4	27.3a.21	27.3.a.20	27.4.a
2016	2 (5)	30 (746)	9 (69)	4 (9)	47 (2058)	3 (4)
2017	4 (18)	23 (600)	11 (64)	3 (38)	49 (1976)	4 (8)
2018	2 (5)	24 (778)	12 (66)	3 (33)	46 (1779)	1 (6)
2019	2 (19)	26 (768)	14 (74)	6 (20)	45 (1788)	4 (7)
2020	5 (66)	25 (790)	14 (62)	6 (96)	42 (1753)	2 (3)

Table: Swedish shrimp landings (in ton) by fishery and ICES subdivision (source: 2016-2020 logbooks).

	SWE_OTB_CRU_32-69_0_0			**SWE_OTB_CRU_32-69_2_22**		
	27.3a.21	27.3.a.20	27.4	27.3a.21	27.3.a.20	27.4.a
2016	1	1015	153	2	784	4
2017	13	72	109	8	630	4
2018	1	716	91	5	472	4
2019	10	532	100	3	369	4
2020	49	627	105	25	379	4

Processing of catch onboard

Most trips sampled involve several fishing operations (hauls) and in each of them fishers sort the catch, extracting fish (and other larger taxa) from Caridea-type and other smaller taxa (Pandalus, Pasiphaea, etc). These "Carideaformes" and smaller taxa are passed through 1-2 sieves with intention of grading the *Pandalus* into two size categories and leaving as remainders the smallest Pandalus and accessory taxa. Observations made during the study period indicate the largest *Pandalus* were retained in the first sieve (here named fraction 91) and landed after being boiled onboard. This first fraction is almost exclusively composed of *Pandalus borealis.* The medium sized *Pandalus* and other medium sized taxa are retained in the second sieve (here named fraction 92) and landed fresh. Laboratory analysis indicate that *Pandalus borealis* constitutes >90% in weight of this fraction, but that a few Pandalus montagui and other species may also be present. The final fraction or remainder (here named fraction 93_94) is either retained and landed fresh, or discarded, and includes a quite variable proportion of *Pandalus borealis* among other smaller taxa (e.g., Pasiphaea).

3. Sampling design and sampling procedures

The sampling design is stratified multi-stage cluster sampling, with vessel as primary sampling unit, trip as secondary sampling unit, haul as tertiary sampling unit, box as quaternary sampling unit and shrimp individual as the quinary sampling unit. The primary sampling units are stratified by quarter and shrimp fishery.

3.1 Vessel selection

Two annual vessel lists (one for each fishery) are built based on previous year's logbook records. Within fishery and quarter, vessels were sampled with probabilities proportional to size without replacement, using the number of trips the vessels made in the fishery during the previous year as a size measure. The sample size per quarter was n=3. During 2016-2018, the function “sample” (base, R CRAN) was used to sample the vessels (the “probs” argument being used to specify vessels initial selection probability). From 2019 to 2020 the function “UPtille” was used (sampling, R CRAN).

3.2 Trip selection

One single trip was sampled from each vessel in each quarter. To arrange the trip, the selected vessels were contacted (by staff using phone in earlier years up to 2016 Q3) or asked to contact the Institute of Marine Research of SLU after receiving a letter (from 2016 Q4 to present). The trip date was decided based on, amongst other, dates of expected vessel activity in the target fishery, weather conditions, and staff and skippers' availability. To reduce refusals, since 2016 Q4, the Swedish authorities have set fines for vessels that do not reply after a series of contact attempts.

3.3. Within-trip sampling methodology and haul selection

Two different methods were used to sample the trips over the study period: In 2016-2017 both fisheries were sampled onboard by scientific observers; In 2018-2020 fishery OTB_CRU_32-69_0_0 kept being sampled onboard by scientific observers but fishery OTB_CRU_32-69_2_22 samples started being collected by fishers themselves (self-sampling). The decision to move towards self-sampling in fishery OTB_CRU_32-69_2_22 was motivated by its low fish discards when compared to fishery OTB_CRU_32-69_0_0 and the need to provide other Swedish fisheries with discard monitoring, namely gill-net fisheries. This change towards self-sampling was accompanied by some changes in the sampling procedures of hauls during these trips:

*Onboard sampling by scientific observers (*OTB_CRU_32-69_0_0: 2016 to 2020[2]; OTB_CRU_32-69_2_22 2016 and 2017)

Two observers were deployed onboard each fishing trip. Sampling of the trip involved, amongst other (e.g., sampling of other species landings and discards):

- In all hauls:
    - Recording the fishing effort (hours trawled)
    - Recording of the weight of the shrimp catch (in kg) by fraction (91, 92, 93_94)
    - Sampling of approximately 1kg (a box) of fraction remainder (93_94)[3]
    - Reference sample of 3 kg of unsorted shrimp (not used in this work)
- In the last haul:
    - Recording the fishing effort (hours trawled)
    - Recording of the weight of the shrimp catch (in kg) by fraction (91, 92, 93_94)
    - Sampling of approximately 1kg (a box) from each fraction (91, 92, 93_94) for further analysis in the lab

*Onboard sampling by fishers (*OTB_CRU_32-69_2_22: 2018 to 2020)

[2] During 2020 the sampling of this fishery was significantly reduced due to the covid pandemic
[3] fraction 94 (also known as “machine discards/remainder” is not always present. This fraction is generally reduced and composed of parts of shrimps and other fish, representing relatively minor weight within the catch. For purposes of present work, its size distribution is assumed similar to the one of fraction 93 (generally known as *lus*) and a combined fraction 93_94 is considered.

Fishers were requested to carry out the sampling onboard:

- In the last haul
    - Record the weight of the catch (in kg) by fraction (91, 92, 93_94) and fate (landed or discarded)
    - Record the size of the sieves they used to separate the shrimp fractions 91, 92 and 93_94
    - Take a 10 kg sample of the unsorted catch (shrimp and fish) before sorting. Provide that sample for further analysis in the lab

3.4 Biological analysis at the lab

All biological samples were kept fresh or frozen before laboratorial analysis. All samples were taxonomically identified, with total weights and total number within each taxa determined. The *Pandalus borealis* in the samples of the last hauls were all individually measured and sexed. Then the samples were stratified in terms of sex, maturity, and parasite status (present/absent) and individual weights determined subject to a maximum of 40 individual's weights per stratum).

4. Data storage and quality checks

The onboard sample data used in the calculation of weights and length composition of *Pandalus borealis* caught by the Swedish fleet was extracted by database Fiskdata2 which is maintained by the Institute of Marine Research of SLU. The database is programmed in Oracle and contains a set of internal quality checking routines that detect some basic errors in data. Onboard sampling data is further checked manually by observers and then semi-automatically by means of r-checks run on sequential data extractions. Fleet level activity and fisheries catch information (e.g., logbook information from which number of vessels, trips, hauls, etc., can be quantified; sales information with insights into processing states of *Pandalus* spp landed, and sold, among other) is annually provided by HaV to SLU upon request. Final FD2 extraction of sampling data used in present analysis took place 2021-12-12.

5. Data preparation for estimation

5.1 Onboard sampling by scientific observers

With regards to trips sampled by scientific observers onboard, data preparation involved the following steps:

a) removal of a few weight-length outliers (with concomitant adjustment of sample weights) using visual inspections of weight-length relationships by area and sex.

b) estimation of species composition of some hauls where 93_94 samples were not collected and some hauls where they were collected but *Pandalus* spp fraction was not identified to species level by means of sequential imputation algorithm.

c) rearrangement of data into three sample-estimation tables

- *Table_catches*: including all catch samples per year*quarter*fleet*fraction*fate, alongside, amongst other, vessel, trip, haul, and box inclusion and selection probabilities.

- *Table_lengths*: including all individuals length measured per year*quarter*fleet*fraction*fate, alongside vessel, trip, haul, and box i inclusion and selection probabilities.
- *Table_length_weight*: including all individuals length and weight measured per year*quarter*fleet*fraction*fate, alongside vessel, trip, haul, box, and individual-weights inclusion and selection probabilities.

5.2. Onboard sampling by fishers

With regards to trips sampled by fishers onboard, data preparation involved the following steps:

a) removal of a few weight-length outliers (with concomitant adjustment of sample weights) using visual inspections of weight-length relationships by area and sex.

b) estimation of length composition of *Pandalus borealis* fractions (91, 92, 93_94) using a sieve model (Daniel Valentinsson, pers. com). In brief the sieve model combines haul-specific sieve information reported by the fishers with the length frequency of Pandalus borealis in the sample to estimate the length compositions of the three size gradings of shrimp.

c) estimation of the proportion in weight of each fraction of shrimp from the estimated length composition (see b) using a standard weight length relationship ($W = 0.001*(L+c)^{2.8951}$ where *c* is an adjustment for length-bin width)

d) assignment of fate (discard/landings) to each fraction – based on fisher's information ….To be detailed

e) rearrangement of the data into three sample-estimation tables

- *Table_catches*: including all catch samples per year*quarter*fleet*fraction*fate, alongside vessel, trip, haul, and box inclusion probabilities.
- *Table_lengths*: including all individuals length measured per year*quarter*fleet*fraction*fate, alongside vessel, trip, haul, and box inclusion probabilities.
- *Table_length_weight*: including all individuals length and weight measured per year*quarter*fleet*fraction*fate, alongside vessel, trip, haul, box, and individual (length,weight) pair conditional inclusion probabilities[4].

6. Design-based estimation

The different sample-estimation tables were used to produce fleet level estimates with the following resolution: year*quarter*fishery*area*fraction (91, 92, 93_94) * fate

[4] note: The (length,weight) pair conditional inclusion probabilities considered the final stratification of sampled individuals into sex*maturity*parasite_status

(landings/discards). The main formulas used are detailed in annex to the present document. In brief, the following methods were used to obtain the estimates:

6.1 Total weight estimates

Inclusion probabilities of samples taken at each stage of the multi-stage design were determined and used in Horvitz-Thompson estimator (Lohr, 2010[i]). This inclusion probabilities contain statistical information of all main elements of the design, including stratification, clustering, sampling method and sample size. In some stages where sampling lists did not exist (e.g., quaternary sampling units – boxes sampled from catch fractions of hauls), total number of possible units (N) was determined using the approximation provided in annex 10 of WKRDB-EST report (ICES, 2019?).

6.2 Length compositions

The original lengths (in 0.01 mm) were rounded to the nearest integer for each record in the sample and the inclusion probabilities of individual length-measured individuals determined. Then the numbers@length were estimated to fleet level using Horvitz-Thompson estimator (e.g., Lohr, 2010). Final estimates were obtained by aggregating numbers@length according to target resolution requested by WKPRAWN 2022.

6.3 Weight-length relationships

Length weight relationships were determined for subdivision 27.3.a.20 samples by year, by year*quarter, and by year*quarter*fishery using a linear model log(weight)~log(a)+b*log(length) where individual observations were weighed by their individual sample weights (using argument “weights” of function “lm” of base R).

Fleet level adjustments

The estimates were aggregated across fleets, fractions, and combinations of subdivisions required to obtain the specific year*quarter*area (3.a; 4*fate (landings/discards) aggregation requested by WKPRAWN 2022.

In brief, data on landings of *Pandalus* per year, area and quarter is gathered from fishery logbooks. The estimates produced from sampling data include numbers@length discriminated as combinations of 3 size categories (91, 92, 93_94) and 2 fate categories (landings and discards). Jointly, 93_94*landed constitutes a low-value fraction that used to be discard but that has become mandatory to land since XXXIn the logbooks 93_94*landed cannot be distinguished from 92 landings also done fresh. Similarly, boiled shrimp and fresh shrimp cannot be distinguished. To match logbooks weights to sample data that distinction [5]

To solve this resolution mismatch, the following procedure was used in present work:

To separate the proportion of “landed boiled shrimp” (91*landed) from “landed raw shrimp” (92*landed) *and* “typical raw landings” (92*landed) from “low-priced landings” (93_94*landed) in logbooks, sales notes were used. Sales notes contain every sale registered at different buyers and, although not constituting an as good a source of information on catches as logbooks (not all trips and catches are accounted for in sales data), they do indicate the weight of the shrimp fraction sold boiled and the weight of the shrimp fraction sold fresh of most trips landing each year. Additionally, they allow the distinction of “typical fresh landings” (92*landed) from “low-priced fresh landings” (93_94*landed)[6] Having identified and quantified the proportions boiled/fresh and typical/low-price fresh landings in sales, those proportions were then applied to logbooks totals to partition logbook data into weight of boiled shrimps, fresh typical shrimps and low-priced shrimps. These weights were used to calculate weight of discard and numbers at length described below

From point estimates of weight of landings and weight of discard per year and quarter in 27.3.a, a discard factor is calculated as weight discard/weight landings. From this factor weight discard can be calculated from true landings described above. Kattegat is seen as very much alike Skagerrak and as there are more samples from Skagerrak discard factor has been borrowed from Skagerrak for each year and quarter. Length distributions in the North Sea are perceived as different from Skagerrak and therefore discard factor has been borrowed from Skagerrak only when there in no discard factor.

Number at length are produced for landings and discards. These numbers were adjusted with true weight for each year, area and quarter.

[5] As boiled shrimp lose weight during boiling and a large proportion of the catches are boiled, logbook records are not useful to derive live weight of shrimp caught y the swedish fishery even if they are broadly considered more reliable than other data sources.
[6] reduction sales can be identified as raw shrimps that are sold to a very low price per kilo, clearly lower than the price for “ordinary” raw shrimp. In the present work a threshold of 15 sek/kg was used.

Filling of missing strata

To be detailed

7. Results

7.1 Implementation

A total of 103 trips and 39 unique vessels were sampled throughout the period. Over the time period most of the sampling took place in the main fishing area of the Swedish fleet, i.e., Skagerrak (27.3.a.20) and only 9 trips were observed with fishing operations in Kattegat (27.3.a.21) and North Sea (27.4.a) (Table XX, Table XX).

A few departures to the initially described sampling design took place over the period. These departures constitute a source of potential bias in the sampling that should try to be fixed / accounted for in future estimation:

- The draw effectively implemented consisted of n=7 vessels from each list. From these vessels, n = 3 were somewhat ad-hoc selected. The reason for such ad-hoc selection (vs. A possible sequential one) is presently unknown. One possibility is that convenience played a role, but the possibility of non-response[7] or even refusals from specific vessels cannot be discarded. The procedure represents a selection bias that is expected to have some (unknown) impact on estimates [needs to be checked]..
- Trips and hauls (in the few instances where they were not censused) were assumed to have been sampled without replacement.
- Vessels from both fisheries might not have been completely independently sampled. A few vessels participate in both fisheries. Part of the reason for adopting without replacement is to reduce the chance of overburdening vessels within a quarter, which might lead to refusals and/or poor data quality. In general, it is expected that high probability within one list will be reflected as low probability of selection in the other list and so it is unlikely that vessels get selected for both fisheries with quarters [needs to be checked].
- In a few vessels were sampled that were not in the original sampling draw. This happened in only a few quarter*fishery combinations. These vessels were for the most selected by convenience, to secure samples when normal sampling failed. For purposes of the present work these vessels were assumed to have been selected in the draw.

[7] Non-responses may range from hard refusals to aspects like lack of observers for the trips and have different levels of impact on estimates. The importance of documenting non-responses has been highlighted by ICES expert groups that have also provided code-lists and guidelines for such documentation (e.g., SGPIDS, 2011, 2012, 2013).

Table 1. Realized sampling in the Swedish fisheries targeting *Pandalus* during 2016-2020: number of vessels sampled. Planned sampling is n=3 per fishery and quarter. Realized samples with different number from planned sampling are in bold.

Year	Quarter	SWE_OTB_CRU_32-69_0_0	SWE_OTB_CRU_32-69_2_22
2016	Q1	3	3
	Q2	3	3
	Q3	3	**4**
	Q4	**2**	3
2017	Q1	3	3
	Q2	3	**4**
	Q3	**2**	3
	Q4	3	**1**
2018	Q1	**2**	3
	Q2	**2**	3
	Q3	3	3
	Q4	**2**	3
2019	Q1	**2**	3
	Q2	3	3
	Q3	3	3
	Q4	**2**	3
2020	Q1	**2**	**4**
	Q2	0	**2**
	Q3	0	3
	Q4	0	3

Table 3. Realized sampling of Swedish *Pandalus* fisheries during 2016-2020

	SWE_OTB_CRU_32-69_0_0					SWE_OTB_CRU_32-69_2_22				
	unique vessels sampled	number of trips sampled	number of hauls sampled	number of hauls sampled for lengths	Number of lengths measure d	unique vessels sampled	number of trips sampled	number of hauls sampled	number of hauls sampled for lengths	Number of lengths measure d
2016	7	11	44	10	6722	10	13	24	13	8701
2017	8	11	43	11	7426	9	11	24	11	8364
2018	6	9	30	9	5697	12	12	12	12	7325
2019	8	10	33	10	6663	12	12	12	12	6541
2020	2	3	6	3	2183	8	12	12	12	7086

Table 4 . Number of trips observed in the Swedish *Pandalus* fishery per subdivision and fishery during 2016-2020

	SWE_OTB_CRU_32-69_0_0					SWE_OTB_CRU_32-69_2_22				
	2016	2017	2018	2019	2020	2016	2017	2018	2019	2020
27.3.a.21 (Kattegat)	0	2	0	0	1	0	0	0	0	0
27.3.a.20 (Skagerrak)	10	6	9	9	1	13	11	12	12	12
27.4.a (North Sea)	1	3	0	1	1	0	0	0	0	0

	SWE_OTB_CRU_32-69_0_0			SWE_OTB_CRU_32-69_2_22		
	27.3a.21	27.3.a.20	27.4	27.3a.21	27.3.a.20	27.4
2016	0 (0)	6 (10)	1 (1)	0 (0)	10 (13)	0 (0)
2017	2 (2)	6 (6)	3 (3)	0 (0)	9 (11)	0 (0)
2018	0 (0)	6 (9)	0 (0)	0 (0)	12 (12)	0 (0)
2019	0 (0)	7 (9)	1 (1)	0 (0)	12 (12)	0 (0)
2020	1 (1)	1 (1)	1 (1)	0 (0)	8 (12)	0 (0)

Table: Unique vessels and trips (in parentheses) by shrimp fishery and ICES subdivision.

	SWE_OTB_CRU_32-69_0_0			SWE_OTB_CRU_32-69_2_22		
	27.3a.21	27.3.a.20	27.4	27.3a.21	27.3.a.20	27.4
2016	0	15412.91	629.52	0	3618.192	0
2017	1540.561	8822.512	5176.19	0	4358.237	0
2018	0	10974.04	0	0	1374.3	0
2019	0	7944.466	1067.469	0	1201.85	0
2020	97.965	669.997	1400.937	0	1793.95	0

Table: Landings (in kg) by shrimp fishery and ICES subdivision (values from sample).

7.2 Design based estimates

Full design-based estimates of discard weights and numbers@length at year*quarter*fleet*fraction*fate resolution are available from authors upon request.

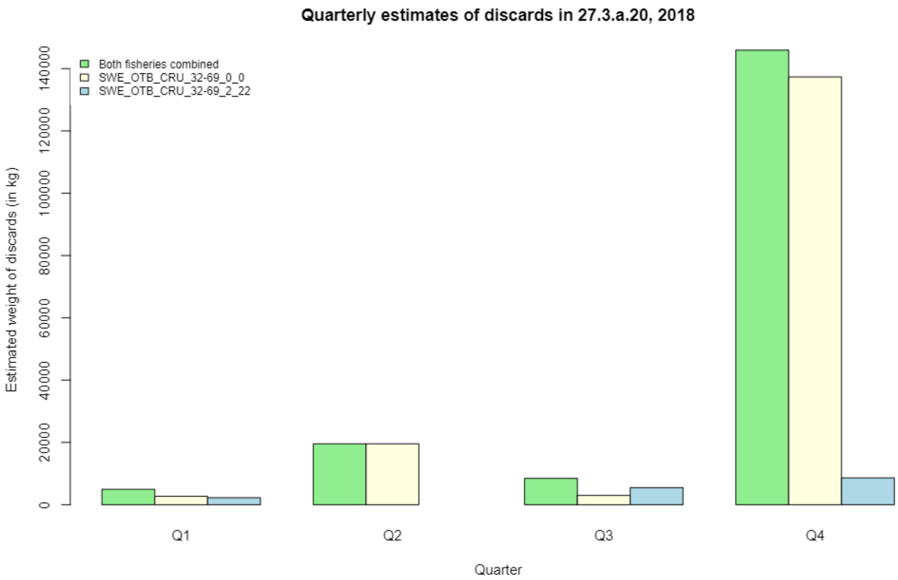


Figure: Quarterly estimates of shrimp discarded by the Swedish fleet operating in ICES 27.3.a.20 during 2018

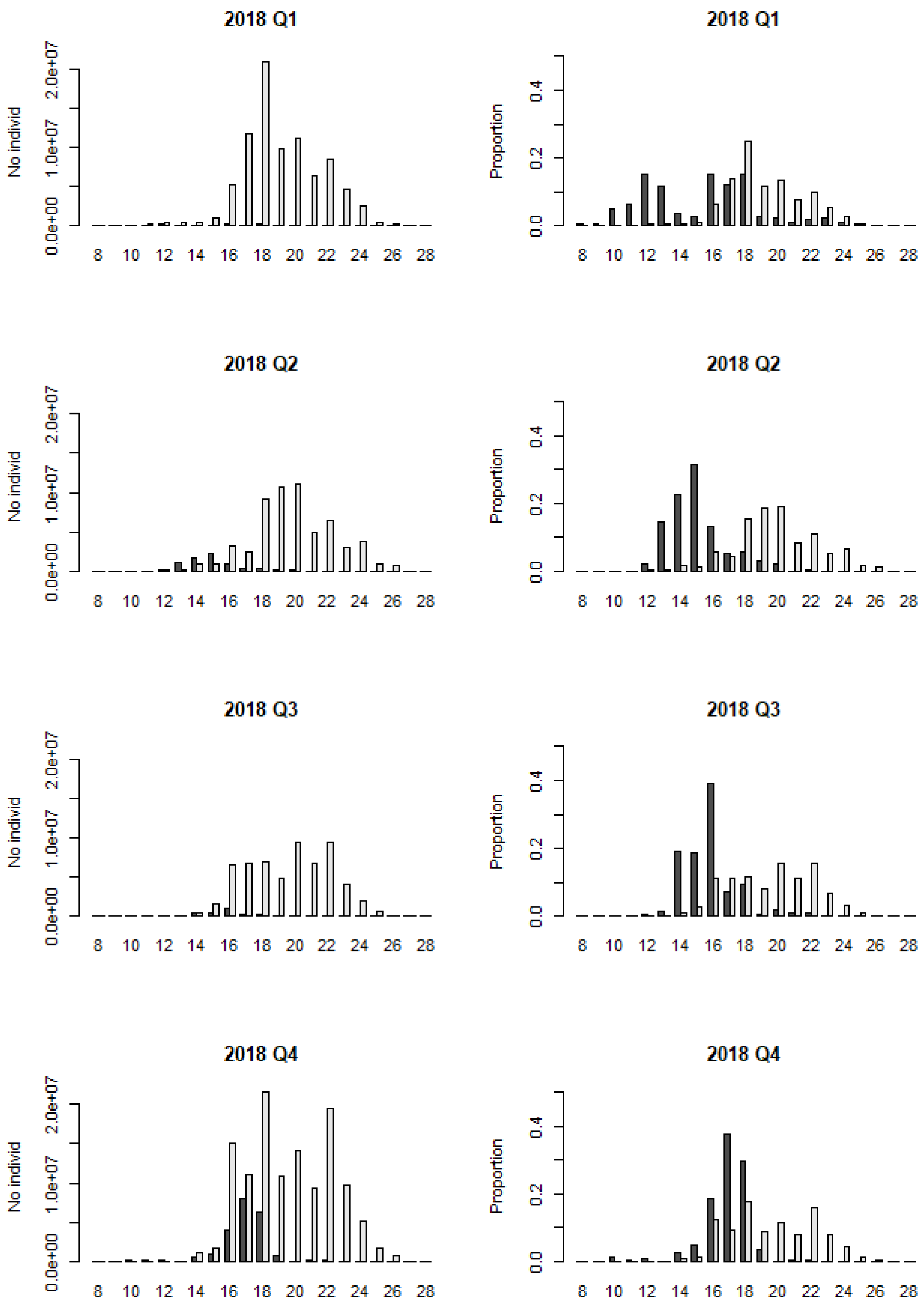


Figure: Quarterly estimates of length distribution of shrimp landed (white bars) and discarded (gray bars) by the Swedish fleet operating in ICES 27.3.a.20 during 2018. Note:

these results might differ from the data submitted to assessment due to different aggregation level and fleet level adjustments to the estimates.

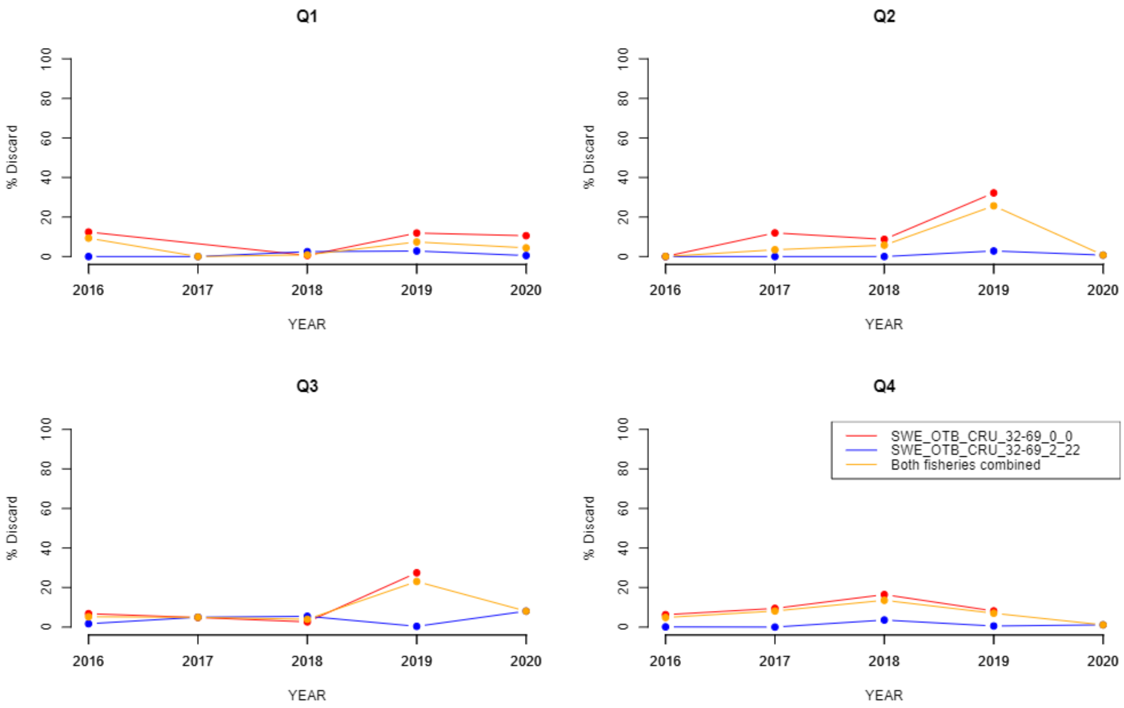


Figure: Quarterly estimates of shrimp discard percentage by the Swedish fleet operating in ICES 27.3.a.20 over 2016-2020.

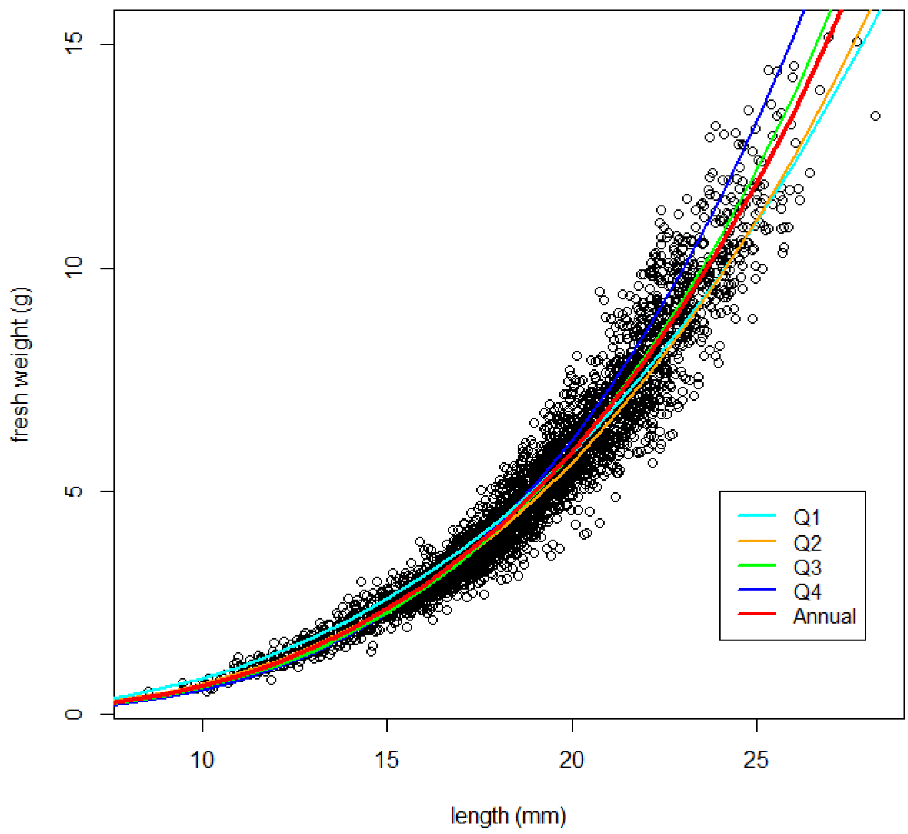


Figure: Quarterly and annual design-based weight-length relationships of shrimp caught in the two Swedish fisheries (combined) during 2018.

7.3. fleet level estimates

Fleet level estimates were provided to NIPAG at requested resolution.

8. Next steps towards improving Swedish estimates of its *Pandalus* fisheries

This WD will be finalized and submitted alongside the WKPRAWN 2022 or NIPAG report 2022. For now, it should should not be cited without prior contact with the authors.

It is envisioned that the estimation methods presented here will be used in upcoming submissions of data to NIPAG, with improvements detailed in future updates or this WD. A few aspects can already be highlighted as worthy of further consideration:

- In the present work the variances of main estimates, namely discard weight, have been preliminary calculated assuming with replacement at PSU level (Lohr, 2010, eq. 6.25). These variances are not yet provided alongside the data because the low sample size originated several data gaps and strata sampled with n=1 (see above) which make computation and display of variances a bit tricky. This aspect and how it can be resolved will be detailed in a future update of this document.

- In the present work, unbiased Horvitz-Thompson estimators were used to estimate both landings (as a intermediary step) and discards. When a good correlation is found between the target variable (i.e., landings or discards) and an auxiliary variable which total value is known from a separate data source (e.g., effort in logbooks) model-assisted ratio estimators provide a (slightly) biased alternative to unbiased Hurvitz-Thompson estimator but that is superior precision. Preliminary analyses of sample data and simulation work carried out in parallel to the current work indicate a reasonable correlation may exist between landings and effort (not so much between discards and effort). That correlation appears to improve precision of estimates even at low (n=3) sample sizes and should be explored in future attempts to improve precision of *Pandalus* estimates.

- A more in-depth look is needed into aspects such as: a) the final fleet level steps (so as to secure a solid basis for subsequent gap filling imputations), b) biases involved in data collected by scientific observations onboard vs data collected by fishers themselves (alongside further exploration of the reference samples from catch that were also collected onboard)

To handle:

Magnus, Lisa, Mikaela, [ask for review]

Can we ask Mats to review? His WD to NIPAG (by email? Are they in his computer?)

It woiuld be great if we can make this WD final at NIPAG (so others can review it and comment) and just available as draft/abstract at WKPRAWN.

1. Acknowledgements

Special thanks to Mats Ulmestrand, Lisa Sörman, Anne Johansson, Kristin Öhman and all onboard observers, lab personnel and programme leaders, involved over the years in data collection of the Swedish Pandalus fisheries.

2. References

Lohr, S.L. (2019). Sampling: Design and Analysis (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780429296284

3. Annex [formulas]

Formulas are derived (adapted? expanded?) from Lohr 2019 Equation (6.25), where the Horvitz-Thompson estimator for two-stage sampling is presented. For the three-stage sampling, a Horvitz-Thompson estimator, which is applied in this working document, the formulas are shown as follows:

$$\hat{t}_{HT} = \sum_{i \in S} \frac{\hat{t}_i}{\pi_i}$$

$$\hat{t}_i = \sum_{j \in S_i} \frac{\hat{t}_{ij}}{\pi_{ij}}$$

$$\hat{t}_{ij} = \sum_{k \in S_{ij}} \frac{t_{ijk}}{\pi_{ijk}}$$

Notations:

- $\hat{t}_{HT}$: *Horwitz-Thompson estimate of the total weights*
- $\hat{t}_i$: *The estimate of the total weights for the i.th vessel*
- $\hat{t}_{ij}$: *The estimate of the total weights for the j.th trip of the i.th vessel*
- t_{ijk}: *The value of the total weights for the k.th haul of the j.th trip of the i.th vessel*
- π_i: *The inclusion probability of the i:th vessel*
- π_{ij}: *The inclusion probability of the j.th trip of the i.th vessel*
- π_{ijk}: *The inclusion probability of the k.th haul of the j.th trip of the i.th vessel*

[i] Lohr, S.L. (2019). Sampling: Design and Analysis (2nd ed.), Page 241 & 245. Chapman and Hall/CRC.

WD 6 Draft. Description of methodology behind compilation of Swedish data on Pandalus borealis (1908-2020) used in *Pandalus* benchmark (WKPRAWN 2022)

By Katja Norén, Swedish University of Agricultural Sciences
Department of Aquatic Resources (SLU Aqua)
Institute of Marine Research

Introduction

Pandalus borealis in 27.4a east (Norwegian Deep) and 27.3a (Kattegat and Skagerrak) has historically been assessed as one stock and length distributions had landings and discards combined. In the benchmark in February 2022 *Pandalus borealis* is however assessed separately for the North Sea and for Skagerrak/Kattegat. Lengths distributions were requested to be produced for landings and discards separately. In summary data was requested per year, quarter and area with length distributions separate for discards and landings.

For 2016-2020 data on discards and lengths is calculated with a new procedure described in the working document (Prista et al., "Sampling design and estimation of commercial catches of northern shrimp *Pandalus borealis* by the Swedish bottom otter trawl fisheries in the 2016-2020").The need for using a new procedure originated from methodologies back in time being insufficiently described and by sudden changes on the SLU staff previously handling this stock~~.~~

For data before 2016, previously submitted values on landings and discards were disaggregated as described in this document. In this document procedures for methods and calculations used back in time to calculate landings, discards and numbers at length are also described.

Material and Methods

In the benchmark on pandalus data over a long time period is used. In this document an attempt is made to describe methods used to create the different parts of this data set. Unfortunately, to the author's knowledge documentation of estimation methodologies back in time is not available or insufficiently described, a situation further aggravated by sudden changes on the SLU staff previously handling this stock. As such the historical compilation on methodologies behind historical data provisions attempted in this document (mainly methodology for discards) should be seen, not as an exact replicate of the procedures then used but rather as a reasonable (good?) description of the methodologies that probably were used during historical data submissions. It must be noted that the basis for the current work consists of actual communication with the data provider for the years (2011, 2012, 2013, 2019) but only for 2019 was the methodology almost fully explained. For other years folders named NIPAG and WGPAND (saved by earlier data providers) were inspected but which files and methodologies were actually used is not known.

The datatypes that are described consists of landings, discards and number at length (ICES 2021).

Results

Landings

Weight of landings used in the stock assessment and modelling is weight of shrimps in fresh weight. Swedish logbooks contain weight of landed shrimp on a detailed level but it does not discriminate between weight of raw shrimp and weight of boiled shrimp. As boiled shrimps lose weight after boiling a factor of 1.13 needs to be multiplied with the weight of boiled shrimp to get fresh weight. Thus it is necessary to compute the weight of landed boiled shrimps to apply the correction factor to this weight.

2016-2020

For the benchmark new calculations were made to produce data for 2016-2020.The amount of boiled and fresh shrimps is derived from Sales notes and the proportion boiled and raw could be calculated per year, area and quarter. From these proportions the weight landed boiled shrimp in the logbook could be calculated and multiplied with 1.13 to get the fresh weight for the boiled fraction. Then total landings in fresh weigh was calculated by summing landed raw weight and landed boiled weight transformed to fresh weight.

1978-2015

To produce data for the benchmark, landings data were taken from the ICES summary sheet and split by quarter and area using estimated yearly proportion from SWE national sources. Landings are corrected for boiling which means that fresh weight is reported.

The description of methodology comes from a thorough description of the year 2019 given by the previous data provider and presented below. This procedure has most probably been the same for a long time but if it is the only procedure is not known. Recalculations of data back in time to check the methodology has not been done. Fresh weight was calculated in several steps. Weight of landed raw shrimp and boiled shrimp presented in reports published by Statistics Sweden (SCB) was used to calculate proportions of boiled and raw shrimps which was used to calculate the weight boiled shrimp from the logbooks (that have landed weight boiled and raw shrimps together). The weight of boiled shrimps was then multiplied with 1.13 to transform weight boiled shrimp into fresh weight. The SCB reports in later years contain total landed weight boiled shrimp and total landed weight fresh shrimps per month with the North Sea and Skagerrak/Kattegat combined. Back in time information on weight boiled shrimps and raw shrimps was probably reported more seldom, perhaps on a yearly basis.

1963-1977

To produce data for the benchmark, landings data was taken from the ICES historical database and split by quarter using average proportion estimated from SWE national sources 1978-1982 (5-years average) and proportion by area as derived from the ICES historical database. Landings are corrected for boiling. The procedure for calculating fresh weigh is not known but probably follow what is described in the paragraph ”1978-2015” above.

1908-1962

To produce data for the benchmark, landings data were taken from the ICES historical database and split by quarter using average proportion estimated from SWE national sources 1978-1982 (5-years average) and proportion by area as derived from the ICES historical database. Landings are not corrected for boiling.

Discard

Swedish discard data has been sampled since 2008. Sampling is mainly conducted in area 27.3.a.20, i.e. Skagerrak. Only very few samples are from 27.3.a.21 i.e. Kattegat or the North Sea. Sampling is done both by observers onboard shrimp vessels and by letting fishermen collect samples from their trip and reporting on weight of landed shrimps and discard (self sampling). A summary of the Swedish sampling strategy for pandalus is presented in table 2 at the end of the Discard section.

2016-2020

For the benchmark new calculations for 2016-2020 were done and these calculations are described in the working document (Prista et al., “Sampling design and estimation of commercial catches of northern shrimp *Pandalus borealis* by the Swedish bottom otter trawl fisheries in the 2016-2020”).

2012-2015

In the shrimp assessments conducted until 2021 data for the North Sea and Skagerrak/Kattegat were combined. For this benchmark yearly data on discards for 2008- 2016 (presented and used in Ices assessments) where allocated using information on logbook landings per year, area and quarter.

The description of the methodology in this section, 2012-2015, is only derived from reading old excel files belonging to earlier data providers and comparing output with values presented in earlier assessments (ICES 2021). Recalculations of data back in time to check the methodology has not been done. It seems the methodology for calculating weight of yearly discard is based on measurements of length distributions of discard samples. A number of trips (shrimp vessels) is selected every quarter. On one haul per trip, samples of lengths were taken from the following catch categories: boiled, raw and discard (lus). In some years also unsorted samples were collected via self sampling. Carapace length was measured on all individuals in the samples and weight was measured until a set number per sex and maturity was weighed.

These lengths were then raised to mirror the whole catch of pandalus. The raising was done in two steps according to the following procedure. In step one, for each trip/haul and catch category, sample weight was calculated using sum of product of numbers at length and weight at length. On each trip/haul the total weight caught of boiled, raw and discard on that trip/haul was also recorded. A trip/haul raising factor step one was calculated as total weight /sample weight per trip/haul and catch category. This raising factor (per trip/haul and catch category) was multiplied with numbers at each length for its specific trip/haul and catch category. In step two, lengths for all trips per quarter were summed for each catch category (boiled, raw and discard). Then these lengths were raised with raising factor step two which was (total weight of landings of pandalus (raw+boiled) per quarter in the Swedish fishery in the North Sea + Skagerrak/Kattegat) / (weight of landed pandalus (raw+boiled) on the sampled trips in the same quarter). This resulted in numbers at length for boiled, raw and discard per quarter. The total weight discard was calculated with sum of product of weight at length and numbers at length.

From trips with staff onboard shrimp vessels not only samples of shrimps to measure were collected. Also weight of landed boiled shrimp and landed raw shrimp was gathered as well as weigh of discarded shrimp. Thus values on discard quota (weight discards/weight landed)* 100 published in assessments of pandalus were derived from these onboard sampling trips.

For 2009 and 2011-2015 there were also samples of unsorted catch from self sampling (fishermen collect samples of unsorted catch). The measured lengths in these samples were partitioned into the catch categories boiled, raw and discard using a sieve model. The sieve model mimics the sieve with a large mesh size used in reality to partition larger shrimps into a high value category that is boiled and the sieve with a smaller mesh size that retains smaller shrimp sold fresh to a lower price. The sieve model was applied separately for each trip and for each trip a specific length-weight model was calculated from the sampled data. As a start value selection range (SR) was set to 2 mm for both sieves and L50 (where half of the individuals are retained) were set to 20 mm (carapace length) for the boiled fraction and 16 mm (carapace length) for the raw fraction. The value for r for both sieves were set to 1.099. Landed weight of boiled and raw shrimp was known and weigh of discard was guessed based on data from other trips. The equation for calculating proportion retained_ is =1/(1+EXP(-r*(Carapace_length(mm) – LD50_for_specific_sieve))). In the able 1 below the equation to calculate proportion retained for each sieve is shown as well calculation of numbers at length for discard, raw and boiled using the sieve models on an unsorted sample.

Table 2. Equations used to partition an unsorted numbers of shrimps into categories: discard, raw and boiled fort different carapace lengths.

Variable to be calculated	Equation

Prop_retained boiled_sieve	=1/(1+EXP(-r * (Carapace_length(mm) – 20))).
Prop_retained raw_sieve	=1/(1+EXP(-r * (Carapace_length(mm) – 16))).
Numbersdiscard	(1- Prop_retained raw_sieve) * Unsorted number at length
Numbers raw	(1- Prop_retained boiled_sieve) * Unsorted number at length-Numbers_lus
Numbers boiled	Prop_retained boiled_sieve * Unsorted numbers at length

As the weight of discard was estimated, the proportion of individuals in the discard fraction (compared to total numbers i.e. discard+raw+boiled) does not correspond to the proportion of weight of discard (compared to total weight i.e. discard+raw+boiled). This was solved using the solver function in excel. L50 values (20 and 16) and SR values 2 and 2 were allowed to vary in order to make the sum of: (percent discard (by weight)-percent discard (by numbers)^2 and (percent raw (by weight)-percent raw (by numbers)^2 and (percent boiled (by weight)-percent boiled (by numbers)^2 add up to zero.

2008-2011

For this benchmark yearly data on discards for 2008- 2016 (presented and used in Ices assessments) where allocated using information on logbook landings per year, area and quarter. Actual procedures for estimating discard during this time period is difficult to conclude even by looking at old excel files from previous data providers. Recalculations of data back in time to check the methodology has not been done. Most probably discard quotas were calculated on onboard trips (dq=weigh discard/(weigh raw + weigh boiled)). Mean discard quotas per quarter were then applied to landings per quarter to get total weigh of discard per quarter. Samples from unsorted catch might have been treated as described in section 2012-2015.

Lengths

Information on lengths is provided by taking samples of shrimps and bringing them back to the lab where length and weight is measured. The sampling strategy is described in table 2 at the end of this document.

2016-2020

For this benchmark new data on lengths were calculated for 2016-2020. Sampling and estimation of lengths is described in the working document (Prista et al., “Sampling design and estimation of commercial catches of northern shrimp *Pandalus borealis* by the Swedish bottom otter trawl fisheries in the 2016-2020”).

2012-2015

For this benchmark a dataset with lengths for area 27.3.a compiled by a previous data provider is used and it probably constitutes of catch i.e. both landings and discards combined. The sampling strategy is described in table 2 at the end of this document. Raising of lengths for: boiled shrimp, raw shrimp and discard is described in paragraph Discards 2011-2015 above. As stated in that paragraph information on the methodology is only derived from reading old excel files belonging to earlier data providers and comparing output with values presented in earlier assessments (ICES 2021). Recalculations of data back in time to check the methodology has not been done.

2005-2011

For this benchmark a dataset with lengths for area 27.3.a compiled by a previous data provider is used. Data for 2005-2007 is probably only landings but unknown. It is also unclear if 2008-2011 is only landings or landings + discard.

The sampling strategy is described in table 2 at the end of this document. The description of the methodology in this section, 2005-2011 is only derived from reading old excel files belonging to earlier data providers and comparing output with values presented in earlier assessments (ICES 2021). Recalculations of data back in time to check the methodology has not been done. For calculating numbers at length per quarter several steps are performed. The number of shrimps per kilo is calculated for raw shrimps and boiled shrimps respectively per quarter. Then total number of shrimps landed (for boiled shrimps and raw shrimps respectively per quarter) is calculated by using landed weight boiled and landed weight raw * numbers per kilo for each catch category and quarter. This total number of shrimps (per catch category, boiled and raw) and quarter is then allocated to the measured numbers at length. The number of measured shrimps per length per quarter and catch category (boiled and raw) is presented as a table with numbers at length presented as percent of column sum. Then a new table with lengths for the same catch categories is constructed and the cell values are calculated with total numbers landed (for each catch category and quarter) * the percentage of shrimps at each specific length for each catch category and quarter.

1990- 2004

For this benchmark lengths for area 27.3.a is used, the data only consists of landings. The sampling strategy is described in table 2 at the end of this document. The procedure for calculating numbers at length is not known.

Summary of Swedish sampling strategy for pandalus 1990-2021

In table 2 below a summary of how trips to sample were selected, what fisheries have been sampled and the method used for sampling is presented. From the sampled trips samples of different catch fractions of shrimp are brought to the lab where for example length and weight is measured. A more thorough description of sampling for 2016-2020 is provided in the working document (Prista et al., “Sampling design and estimation of commercial catches of northern shrimp *Pandalus borealis* by the Swedish bottom otter trawl fisheries in the 2016-2020”).

Table 3. Summary of Swedish sampling strategy for pandalus 1990-2021.

Year	Method for vessel selection (trips)	Fisheries sampled	Metod for sampling
2021	Unequal probability sampling	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid and tunnel. Self sampling (fishermen samples and reports info on landings and discards) on vessels with grid and no tunnel.
2020	Unequal probability sampling	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid and tunnel. Self sampling (fishermen samples and reports info on landings and discards) on vessels with grid and no tunnel.
2019	Unequal probability sampling	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid and tunnel. Self sampling (fishermen samples and reports info on landings and discards) on vessels with grid and no tunnel.
2018	Unequal probability sampling	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid and tunnel. Self sampling (fishermen samples and reports info on landings and discards) on vessels with grid and no tunnel.

2017	Unequal probability sampling	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid without tunnel and on vessels with grid and tunnel.
2016	Unequal probability sampling	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid without tunnel and on vessels with grid and tunnel.
2015	Unequal probability sampling	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid and tunnel. Unclear if sampling on vessel with grid and no tunnel was done by staff or self sampling (fishermen samples and reports info on landings and discards).
2014	Unequal probability sampling	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid and tunnel. Self sampling (fishermen samples and reports info on landings and discards) on vessels with grid and no tunnel.
2013	Unequal probability sampling	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid and tunnel. Unclear if sampling on vessel with grid and no tunnel was done by staff or self sampling (fishermen samples and reports info on landings and discards).
2012	Unequal prob onboard + self sampling [2 foCatNat]	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid and tunnel. Unclear if sampling on vessel with grid and no tunnel was done by staff or self sampling (fishermen samples and reports info on landings and discards).
2011	Unequal prob onboard + self sampling [2 foCatNat]	Vessels with grid and vessels with grid + tunnel	Onboard sampling by staff on vessels with grid and tunnel. Unclear if sampling on vessel with grid and no tunnel was done by staff or self sampling (fishermen samples and reports info on landings and discards).
2010	Unequal prob onboard + self sampling [2 foCatNat]?	Unclear if sampling is done on both vessels with grid and vessels with grid + tunnel	Unclear if sampling was done by staff or self sampling (fishermen samples and reports info on landings and discards).
2009	Ad hoc?	Ad hoc?	Onboard sampling by staff
2008	Ad hoc?	Ad hoc?	Onboard sampling by staff
2007	Unclear	Unclear	Onboard sampling by staff
2006	Unclear	Unclear	Onboard sampling by staff
2005	Unclear	Unclear	Onboard sampling by staff

1990-2004		Samples were taken from the first boat willing to provide samples.	Samples were collected by Fiskeauktionen Smögen after order from the Marine lab. Samples were collected from the first boat willing to provide samples.

References

ICES (2021). Northern shrimp (Pandalus borealis) in divisions 3.a and 4.a East (Skagerrak and Kattegat and northern North Sea in the Norwegian Deep). Report of the ICES Advisory Committee, 2021. ICES Advice, pra.27.3a4a.

WD 7 Ensemble Model of Northern shrimp (Pandalus borealis) in ICES division 3a and 4a east

By Francesco Masnadi, Massimiliano Cardinale, Alessandro Orio, Christopher Griffiths and Mikaela Bergenius Nord

The majority of the contents of this working document are included in section 3.4.11

WD 8 Shrimp in the Barents Sea - assessment models

By F.Zimmermann, C.Hvingel

The majority of the contents of this working document are included in section 4.9, 4.10, 4.11.

WD 9 Estimating size at sex change northern shrimp (*Pandalus borealis*).

By Mikaela Bergenius Nord, Patrik Börjesson, Guldborg Søvik.

Summary

Yearly estimates of size at sex change (L50) for northern shrimp (*Pandalus borealis*) in Skagerrak were computed from individual data (length and sex), collected as part of the annual Norwegian shrimp survey and three different sampling programs of Swedish commercial catches (onboard sampling and/or harbour sampling of unsorted and sorted catches). The current estimate of L50 that is being used in the ICES stock assessment, which is constant with time, is based on part of the Swedish commercial data. The aim of this study was to investigate if L50 has varied over time, and the sensitivity of the L50 estimate to datatype (fishery dependent and independent) and the timing of the data collection. A time variant size at sex change will likely influence the productivity of the stock and its dynamic, and is therefore important to consider in stock assessment, the associated computation of reference points and future fishing opportunities.

The analyses show that the estimates of L50 varies over time. L50 computed from the Norwegian survey (1984-2019) varied between a minimum of 16.3 mm (+/- confidence interval (CI) 0.4 mm) and maximum value of 21.8 mm (+/- CI 0.4 mm). L50 computed from the combined Swedish fishery dependent sampling programs varied between a minimum of 16.1 mm (+/- CI 0.8 mm) and a maximum of 19.9 mm (+/- CI 1.2 mm). The yearly estimates of the parameter for the different data types show in general the same temporal pattern, with the exception of three deviating years. Also, the absolute values of L50 were in general comparable, with the exception of the same three years, during which the survey showed distinctly higher parameter values. The uncertainty (95 percent CI) around the L50 were higher for the fishery dependent data, and can likely be explained by the lower number and less geographic spread of the samples. The effects of this time variant L50 parameter should be evaluated in the stock assessment.

As the shrimp change sex in the beginning of the year (quarter 1) and the recruits reach a size at which they are selected in the fishing gear during the second half of the year, the timing of the sampling will influence the estimate of L50. The comparatively large number of samples and geographical coverage of the survey make the estimates of L50 from this data source the likely most suitable for inclusion in the stock assessment. Decisions on how to proceed will be taken at the data workshop of *Pandalus* in November 2021.

Introduction

In the Northeast Atlantic, northern shrimp (*Pandalus borealis)* range in its distribution from the North Sea and Skagerrak to the north of Svalbard, and are generally found at depths from 50 to 500 meters at low temperatures (0-5°C) and high salinities (34.1-35.7 ‰) (Shumway *et al.*, 1985). In the Northeast Atlantic, two commercially important stocks of *P. borealis* exist, the Barents Sea stock and the Norwegian Deep/Skagerrak (NDSK) stock.

The species is a protandric hermaphrodite, where each individual first matures and reproduces as male, followed by a transitional or intersexual phase before becoming female (Shumway *et al.*, 1985). However, in the southern part of the distributional range, several authors have reported early maturing females, so called primary females (Shumway *et al.*, 1985). Allen (1959) wrote that in the North Sea, more than 30 % may never show male characters. The female shrimps spawn once a year, extruding the eggs in late summer to early autumn before carrying them on the pleopods until hatching commences in spring the following year. The size (and age) of sex change for the pandalid shrimp is likely not a fixed phenotypc trait (Charnov and Anderson, 1989) but likely a response to a combination of factors, such as population density, size frequency distribution, adult mortality rate, and resource availability (Charnow, 1982, Baeza, 2007). A variation in the size at sex change is likely to influence shrimp stock productivity and dynamics, complicated by the fact that, as for many other marine species, reproductive output increases with an increasing female size (Parsons and Tucker, 1986).

In the present study, we investigate the variation in size at sex change (L50) over time for the northern shrimp stock in NDSK. In the current stock assessment, a fixed value for size at (female) maturity is used over the entire time series (ICES, 2016). This value was derived from Swedish commercial catch samples from 2013-2014. Size at maturity (for females) will be roughly the same as size of sex change, if we assume that the ovaries start developing immediately after the female stage is reached (following 2-3 intersex moults). As well as studying the suitability of a constant L50, we investigated the appropriateness of using Swedish commercial catch samples for its estimation, as is done now (ICES, 2016), as opposed to using data from the Norwegian shrimp survey. This survey has been conducted annually since 1984 (Søvik and Thangstad, 2020), during which a large number of *P. borealis* has been sexed and length measured. The survey also has a substantially larger spatial cover in comparison to the Swedish commercial data. The survey has some limitations, however, including that the timing of the expedition has changed over time. As *P. borealis* naturally grows throughout the year, but also changes sex from male to female, the estimate of L50 may be influenced by the time of sampling. A second limitation of the survey data is that the sex/stage of the 1-year old shrimps has not been determined; they have all been assumed to be males which is not correct.

Thus, with this study our objective is to confirm or reject the current fixed value of L50 (17.87 mm) (female maturity), and in the case of rejection, recommend the most appropriate data to be used for its estimation.

Material and Methods

Data

In this section, we first describe the data derived from the Swedish commercial fishery used to estimate the size at sex change, i.e. the carapace length at which 50 percent of the *Pandalus* shrimp individuals were females (L50; 3.1.1). We then describe the Norwegian shrimp survey data (3.1.2).

Swedish fishery dependent data

Data on shrimp lengths and sex have been collected routinely from the Swedish shrimp fishery since 1990 as part of three different sampling programs. Data from the sampling program 13 are available for the years 1990 -2004. The programs 11 and 22 started in 2010 and are ongoing programs today. Data for the years 2005-2009 are available, but not quality checked and entered into the database, and are therefore not included in this analysis. The data were collected either on-board by observers from SLU Aqua or at the fisheries auction, or by the fishers themselves as part of a self-sampling program, for which the samples were either collected by SLU Aqua from the fishers in the harbours, or sent directly to the lab. At the lab, all shrimp were sexed and the carapace length measured. Vessels were selected for sampling by a random draw (with replacement) from a list of operating vessels with unequal probability (the probability was proportional to the number of trips). Here, only data from Skagerrak were included, and this is also where the vast majority of the samples have been taken. See Annex 1 for a brief description of the data from the different sampling programs, data preparation and a sensitivity analysis of combining data from the different programs into one time series.

Norwegian shrimp survey

The Norwegian Bottom Trawl Survey for northern shrimp (*Pandalus borealis*) in the Skagerrak and Norwegian Deep (ICES Divs. 27.3.a and 27.4.a. east) has been conducted by the Norwegian Institute of Marine Research (IMR) since 1984 (Søvik and Thangstad, 2020). Until 2003, the survey was conducted in quarter 4. In 2004 and 2005, the timing of the survey was changed to May-June, and in 2006 the survey period was again changed, to the 1st quarter (January/February) to improve the estimates of spawning stock biomass (SSB) (berried females) and recruitment (1-year old shrimp which have not yet entered the fishery). In addition to data on shrimp lengths and sex, data on multiple demersal fish species preying on the northern shrimp are collected. The survey is based on a stratified survey design with fixed trawl stations, covering depths from 100–550 meters in four depth zones (100-200m, 200-300m, 300-500m and >500m, see Figure 1). The stratification has been revised over time and the number of strata has been reduced, but the depth zones have been maintained. The list of sampling stations has also been revised and currently consists of 111 fixed stations, of which 60 stations are located in the Skagerrak. Only data from Skagerrak were included in this analysis, to be comparable to the Swedish commercial data.

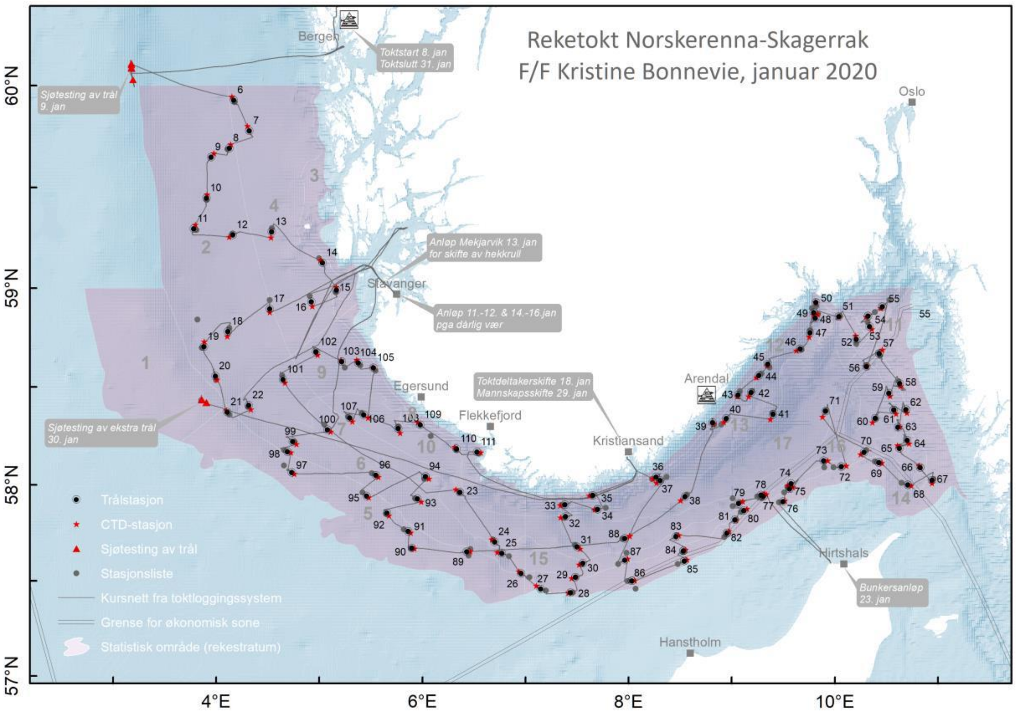


Figure 1. The Norwegian shrimp survey in ICES Divs. 3.a and 4.a east (Skagerrak and the Norwegian Deep) in January 2020 with R/V Kristine Bonnevie: sailing route with fixed stations (grey dots), testing area for trawl gear (red triangles), trawled stations (black dots), and CTD-stations (red stars). (Map from Søvik and Thangstad, 2020).

The Norwegian shrimp survey uses a 'Campelen 1800' trawl with 14-inch rockhopper gear. The mesh size of the codend is 20 mm, and a 10 mm mesh size inner net is used. From 2008 and onwards strapping has been used to keep gear geometry constant, with an intended door spread of 48-52 meters irrespective of depth. With a trawling duration of 30 minutes and a speed of 3 knots, the swept area will be approximately 0.11 km^2.

Data analysis

Shrimp length and sex data collected as part of the three different fishery dependent sampling programs were combined into one dataset (Annex 1). By combining the data from the different programs, we make the assumption that the differences in sampling within the different programs do not result in marked differences in L50 over time (Annex 1). Yearly estimates of the size at sex change, specified as the size at which 50 percent of the individuals were females (L50), were subsequently computed from both the combined fishery dependent data and the Norwegian survey data, so that these could be compared. Because shrimp grow throughout the year, the timing of sampling can influence the estimate of L50. The Norwegian survey has since 2006 been conducted in quarter 1, to give good estimates of the 1-group (recruitment) and the spawning population (berried females) which will hatch their larvae in March-April the same year (ICES, 2005). L50 estimated from fishery dependent data were therefore in the primary analysis limited to samples collected in quarter 1, to be comparable with the L50 computed from the survey data.

Size at maturity (L50) (for females) is the input value required for the stock assessment model. Mature females are defined to be all second-time spawners in the stock as well as the first-time spawners that will contribute to stock productivity in the year for which quota advice is given. In quarter 4, all females are considered to contribute to the following year's productivity (= advice year), i.e. first-time spawners with head roe in samples are assumed to mate and put out roe within weeks. Shrimp change sex from male to female in winter-spring, and the transitional or intersex stage occurs in the population mainly from January to April. The intersex shrimp in quarter 1 will spawn and mate as females half a year later, in autumn, thus contributing to the production the following year. The same applies to the relatively few females with no roe present in samples in quarter 1. Hence, only berried females and females that had just hatched their eggs were included as (mature) females in the L50 estimate from quarter 1 for both the survey and the commercial data. These females contribute to the stock production the same year as the sampling takes place (= advice year). Any second-time spawner with no roe, but which is definitely mature, was included as well. In quarter 2, all female stages, as well as the intersex individuals, are considered to contribute to the following year's production (= advice year), thus they are all considered (mature) females.

As the Norwegian survey was conducted in quarter 4 from 1984 to 2003, and in quarter 2 in 2004 and 2005, L50 from the fishery dependent data were in two additional scenarios computed only including data from respectively quarter 4 and quarter 2. As the number (and thus geographical distribution) of samples are substantially reduced when limiting samples to one quarter only (Annex 1), yearly estimates of L50 were also computed based on the entire commercial data set, including all quarters. Another factor that contributes to a difference in the absolute estimate of L50 between the survey and the commercial data are the different fishing gear used. The survey trawl (10 mm mesh size in inner net in cod end) selects for smaller individuals than the commercial gear (minimum legal mesh size is 35 mm, but mesh size in the Swedish fleet has increased lately). Furthermore, as already mentioned, the 1-year old shrimp caught on the survey are not sexed, but just taken to be males. All specimens in the Swedish commercial samples have been sexed, but the youngest age group makes up a small proportion of the Swedish data. While the absolute values of L50 from these different data types therefore are not expected to be the same, the temporal trends of L50 could be similar.

L50 was estimated fitting a generalized logistic mixed model, with year as a fixed effect and haul as a random effect, via maximum likelihood, to the proportions of females at carapace length:

$$p = \frac{\mathrm{I}}{1+e^{-k(CL-L_{50})}}$$

where CL is the carapace length in mm, K is the slope of the logistic curve and L50 is the carapace length at which 50 percent of the shrimp numbers were females (that is, the size at sex change). Figure 2 illustrates an example of the logistic function describing the proportion of females at carapace length (i.e. an ogive of sex change).

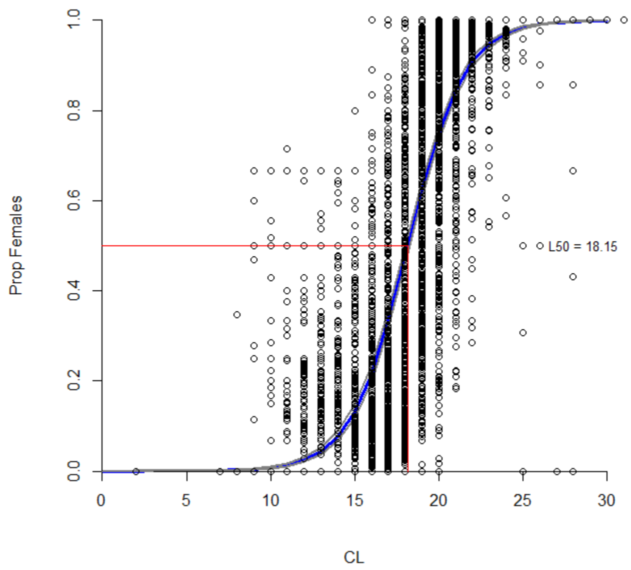


Figure 2. An example of an ogive of sex change for *Pandalus borealis* where CL refers to carapace length (in mm) and Prop female to proportion of females. L50 refers to the length at which 50% of the individuals are females.

Results

Size at sex change (L50) of northern shrimp, *Pandalus borealis,* in Skagerrak was variable over the years (Table 2, Figure 3). The overall pattern between years was relatively similar irrespective of data type (survey or fishery dependent), with the exception of some years (2004 and 2014, in particular) when the survey estimates show a clear increase compared to previous years. The yearly estimates of L50 were also similar in absolute terms, with the estimated uncertainty ranges (+/- 95 percent confidence intervals) overlapping for the survey and the commercial data from quarter 1 (Table 1, Figure 3), again with the exception of some years. The 95 percent confidence intervals of the estimates from the fishery dependent data were however, much larger than those of the survey estimates, most likely due to the substantially lower number of samples, and thus spatial cover (Annex 1). Furthermore, values of L50 in absolute terms computed from the quarter 1 commercial data were also similar to the survey estimates for the years 1984-2002, even though the survey in this time period was conducted in quarter 4, and the shrimps by then would have had time to grow to a larger L50. As described in the methods section however, the data from the Norwegian survey do include the small young (0 year old) individuals, not captured in the commercial gear (Figure 4), making the comparison of absolute values complicated.

Figure 5 illustrates the yearly estimates of L50 computed from fishery dependent samples collected in quarter 2 and 3, as well as from all quarters combined. The figure reveals no obvious improved agreement between estimated L50 from the survey and commercial data, when the latter was based on samples from either quarter 2 or 4 and compared to the L50 from the equivalent survey period, (quarter 4 for the period 1984-2002, and quarter 2 for 2004 and 2005).

Table 1. Maximum (L50max) and minimum (L50min) size at sex change (L50) of northern shrimp, *Pandalus borealis*, in Skagerrak, estimated from the samples collected as part of the Norwegian survey and Swedish fishery dependent data from quarter 1, quarter 2, quarter 3, and all four quarters. The +/- 95% CI refers to the upper and lower 95 percent confidence intervals.

Scenario	L50min (mm +/- 95% CI)	L50max (mm +/- 95% CI)	Years
Norwegian survey	16.3 (16.0, 16.7)	21.8 (21.3, 22.2)	1984-2019
Fishery dependent data from quarter 1	16.1 (15.0, 16.9)	19.9 (19.4, 21.2)	1990-2019
Fishery dependent data from quarter 2	14.2 (9.5, 16.2)	18.4 (17.8, 19.0)	1990-2019
Fishery dependent data from quarter 3	17.3 (14.2, 18.0)	21.2 (20.4, 22.1)	1990-2019
Fishery dependent data from all quarters	16.9 (16.4, 17.4)	19.7 (19.2, 20.4)	1990-2019

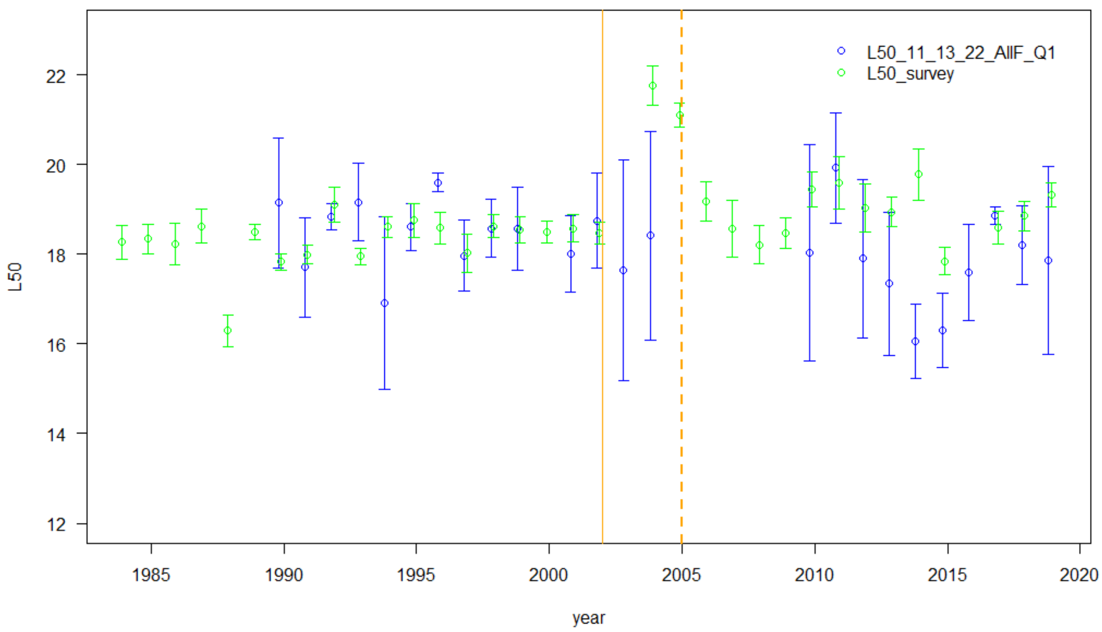


Figure 3. Size at sex change (L50) of northern shrimp, *Pandalus borealis*, in Skagerrak, estimated from the samples collected as part of the Norwegian survey (L50_survey) and Swedish fishery dependent data from quarter 1 (L50_11_13_22_AllF_Q1). The yellow vertical lines indicate when the timing of the Norwegian survey was changed. Between 1984-2002 it was conducted in quarter 4; in 2004 and 2005 it was conducted in May-June (quarter 2) and since 2006 it has been conducted in quarter 1. The error bars around the point estimates of L50 are the upper and lower 95 percent confidence intervals.

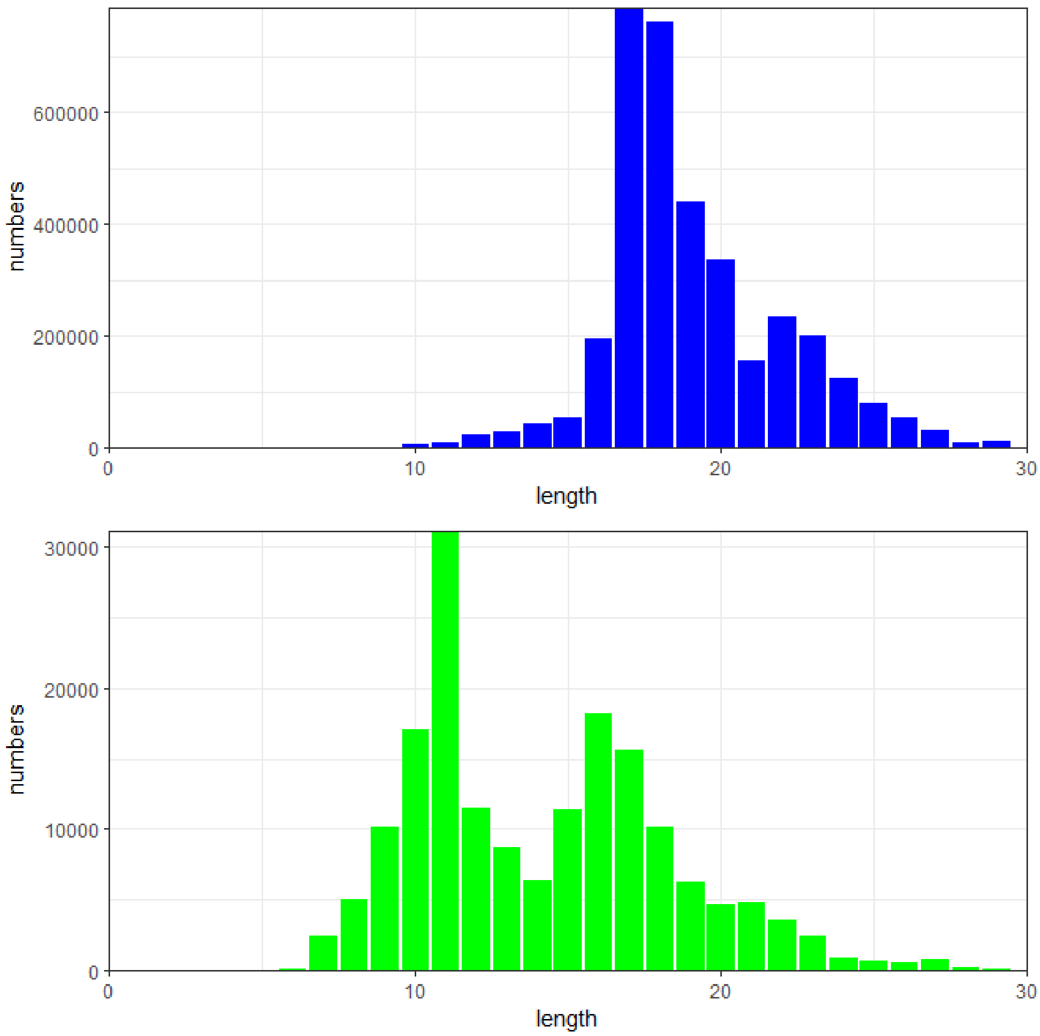


Figure 4. Size frequency distribution of northern shrimp, *Pandalus borealis*, in Skagerrak, estimated from the samples collected as part of the Norwegian survey (1984-2019, bottom panel) and Swedish fishery dependent data from quarter 1 (1990-2019, top panel). Note that the length samples in the Swedish fishery dependent data have been raised to the total catches in this figure, thereof the large numbers.

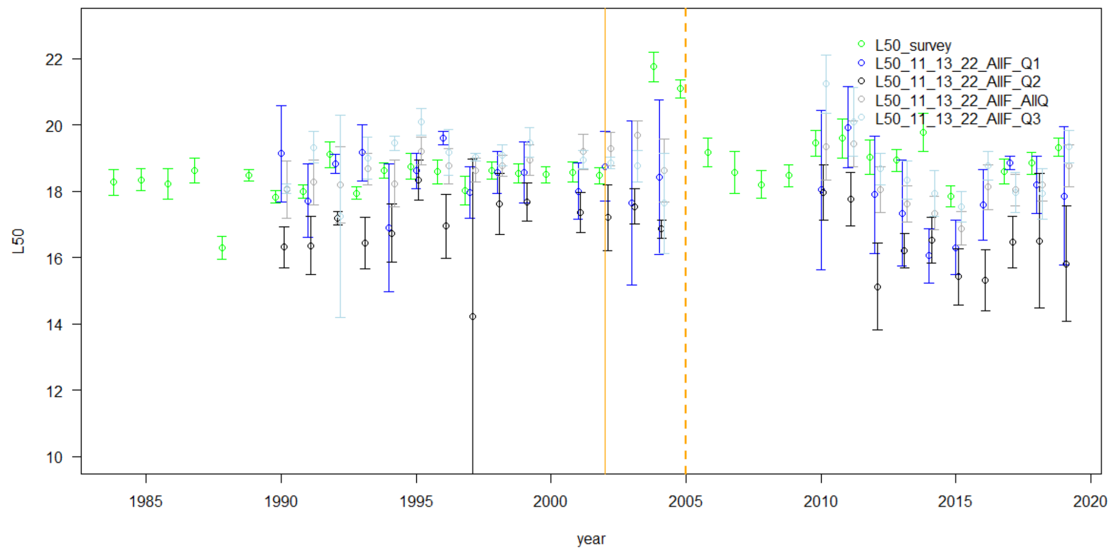


Figure 5. Size at sex change (L50) of northern shrimp, *Pandalus borealis*, in Skagerrak, estimated from the samples collected as part of the Norwegian survey (L50_survey) and Swedish fishery dependent data from quarter 1 (L50_11_13_22_AllF_Q1), quarter 2 (L50_11_13_22_AllF_Q2), quarter 3 (L50_11_13_22_AllF_Q3), and all four quarters (L50_11_13_22_AllF_AllQ). The yellow vertical lines indicate when the timing of the Norwegian survey was changed. Between 1984-2002 it was conducted in quarter 4; in 2004 and 2005 it was conducted in May/June (quarter 2) and since 2006 it has been conducted in quarter 1. The error bars around the point estimates of L50 are the upper and lower 95 percent confidence intervals.

Discussion

This study shows that there is some annual variability in size at sex change (L50) of the northern shrimp, *Pandalus borealis*, in Skagerrak, since the mid-1980s (with the reservation for a lack in statistical testing). L50 estimated from the survey displayed a somewhat larger range of values (span between max and min) than those computed from the commercial data in quarter 1. The uncertainty around the average estimates from the commercial data however, were substantially larger. The data evaluation and results from this study indicate that the fixed value of L50 of 17.87 mm currently used in the assessment of the northern shrimp stock in NDSK should be evaluated. A time-varying value of L50 may be included in the stock assessment model to investigate its implication for the perception of stock productivity and reference points. The results of this study point to that the survey would be the most appropriate data source to use for the estimation of L50. The survey has a larger number of samples and spatial cover than the Swedish commercial data in quarter 1. Hoverer, the small recruits in the survey data have not been sexed. The proportion of berried females of the 2-year age group in the survey indicates that between 10 and 20% of the 1-year old shrimp are females, i.e. primary females. Taking this into account, means that the line in Figure 1 would be lifted from 0, altering the slope of the logistic curve (K), but likely not the L50-value. There are also other potentially biasing factors when using both survey and commercial data. For example, the survey has been conducted over three different time periods, although with the same gear but with different vessels. The different L50 values estimated from the commercial data from different quarters indicate that the timing of sampling may have a significant effect on the absolute value. Figure 5 shows that it is especially the L50 values from quarter 2 that are different. This makes sense as only male shrimp grow over the winter (from quarter 4 to quarter 1) while berried females don't moult and therefore cannot grow. On the other hand, the youngest age group grows and will be selected more by the survey gear in January compared with in October, possible shifting the sex ratio. Anyhow, demographic changes from quarter 1 to quarter 2 are likely larger, with all shrimps moulting and growing, and the largest females dying and disappearing from the population.

The commercial data have issues with changing sampling programs, spatial coverage and a limited number of samples, especially if quarter 1 is to be used in isolation. The results from this study will be discussed with shrimp experts at the ICES Pandalus working group in February and September, for decisions on how to proceed.

This study, among several others, confirm that the size at sex change in shrimp is plastic, and can change over time (Charnov and Anderson, 1989, Koeller *et al*., 2000) and/or between locations (Bergström, 1992). Many factors have been suggested to trigger the onset change (at a different size or age) such as, environmental factors (Wieland, 2004), changes in stock size (Jonsdottir *et al*., 2017), density dependent growth (Koeller *et al*., 2000) and evolutionary effects (Charnov and Skuladottir, 2000), but the overall conclusion seems to be that there are several factors working effecting the change. A next step in this analysis could be to investigate the potential drivers of the change in size at sex change of *Pandalus borealis*, for instance taking into account the considerable reduced stock size since 2011-2012 (ICES, 2021). Since the new assessment model being proposed for the benchmark, the potential spatial variability of this parameter across Skagerrak, and between Skagerrak and the Norwegian Deep should also be investigated.

References

Allen, J. A. 1959. On the biology of *Pandalus borealis* Krøyer, with reference to a population off the Northumberland coast. Journal of the Marine Biological Association of the United Kingdom, 38: 189-220.

Baeza, J. A. 2007. Sex allocation in a simultaneously hermaphroditic marine shrimp. Evolution, 61: 2360–2373.

Bergström, B. 1992. Growth, growth modelling and age determination of *Pandalus borealis*. Marine Ecology Progress Series, 83: 167–183.

Charnov, E. L. 1982. The Theory of Sex Allocation. Princeton University Press, Princeton, NJ.

Charnov, E. L., and Skuladottir, U. 2000. Dimensionless invariants for the optimal size (age) of sex change. Evolutionary Ecology Research, 2: 1067–1071.

ICES, 2021. ICES Advice 2021 – pra.27.3a4a – https://doi.org/10.17895/ices.advice.7834

ICES 2016. Report of the Benchmark Workshop on *Pandalus borealis* in Skagerrak and Norwegian Deep Sea (WKPAND) 20–22 January 2016, Bergen, Norway. ICES CM 2016/ACOM:39. 71 pp.

ICES 2005. Report of the Pandalus assessment working group, 27 October – 5 November 2004. ICES C.M. 2005/ACFM:05, 74 p.

Jónsdóttir, IG (2017) Predation on northern shrimp (*Pandalus borealis*) by three gadoid species, Marine Biology Research, 13:4, 447-455,

Jo´nsdo´ttir, I. G., Tho´rarinsdo´ttir, G. G. and Jonasson, J. P (2017). Influence of decreased biomass on the ogive of sex change of northern shrimp (*Pandalus borealis*). – ICES Journal of Marine Science, 75: 1054–1062.

Koeller, P. A., Mohn, R., and Etter, M. 2000. Density dependant sex change in northern shrimp, *Pandalus borealis*, on the Scotian shelf. Journal of Northwest Atlantic Fishery Science, 27: 107–118.

Parsons, D. G., and Tucker, G. E. 1986. Fecundity of northern shrimp, *Pandalus borealis*, (Crustacea, Decapoda) in areas of the Northwest Atlantic. Fishery Bulletin, 84: 549–558.

Søvik, G. and Thangstad, T. H. 2020. Results of the Norwegian Bottom Trawl Survey for Northern Shrimp (*Pandalus borealis*) in Skagerrak and the Norwegian Deep (ICES Divisions 3.a and 4.a East) in 2020. NAFO/ICES PANDALUS ASSESSMENT GROUP MEETING – FEBRUARY AND OCTOBER/NOVEMBER 2020. Serial No. N7032 SCR Doc. 20/001

Wieland, K. 2004. Length at sex transition in northern shrimp (*Pandalus borealis*) off West Greenland in relation to changes in temperature and stock size. Fisheries Research, 69: 49–56.

Annex 1

This annex presents a brief description of the Swedish commercial data used to estimate the size at sex change (L50) of *Pandalus borealis* in Skagerrak. It also presents a sensitivity analysis of combining data from the different sampling programs. Table 1 presents three different sampling programs undertaken by SLU Aqua between 1990 until present, at which individuals of the shrimp *Pandalus borealis* have been collected, sex determined and measured.

Table A1. Sampling program, strategy, sampled fractions, start and end year of the sampling programs included in this analyses. The Strategy refers to if the samples were taken on-board the vessel by SLU Aqua personnel (Observers), personnel at the fish auction in Gothenburg (Observers*) or by the fishers themselves (Self-sampling) and collected by or, sent to, the SLU Aqua lab. Sampled fractions refer to the type of the catch which were sampled: larger shrimp not passing through any of the sorting grids and are boiled for direct consumption (boiled), shrimp that has passed through the first sorting grid and is sold as raw (raw) and small shrimp that goes through both grids in the sorter (lus). For sampling 11 a sample of 0.5 kg lus, 1 kg raw and 2 kg boiled, were taken and analysed in the lab. For sampling 22, a 3 kg sample was taken and for 13 1 kg raw and 2 kg boiled. The Start year and End year indicate when the programs started and ended.

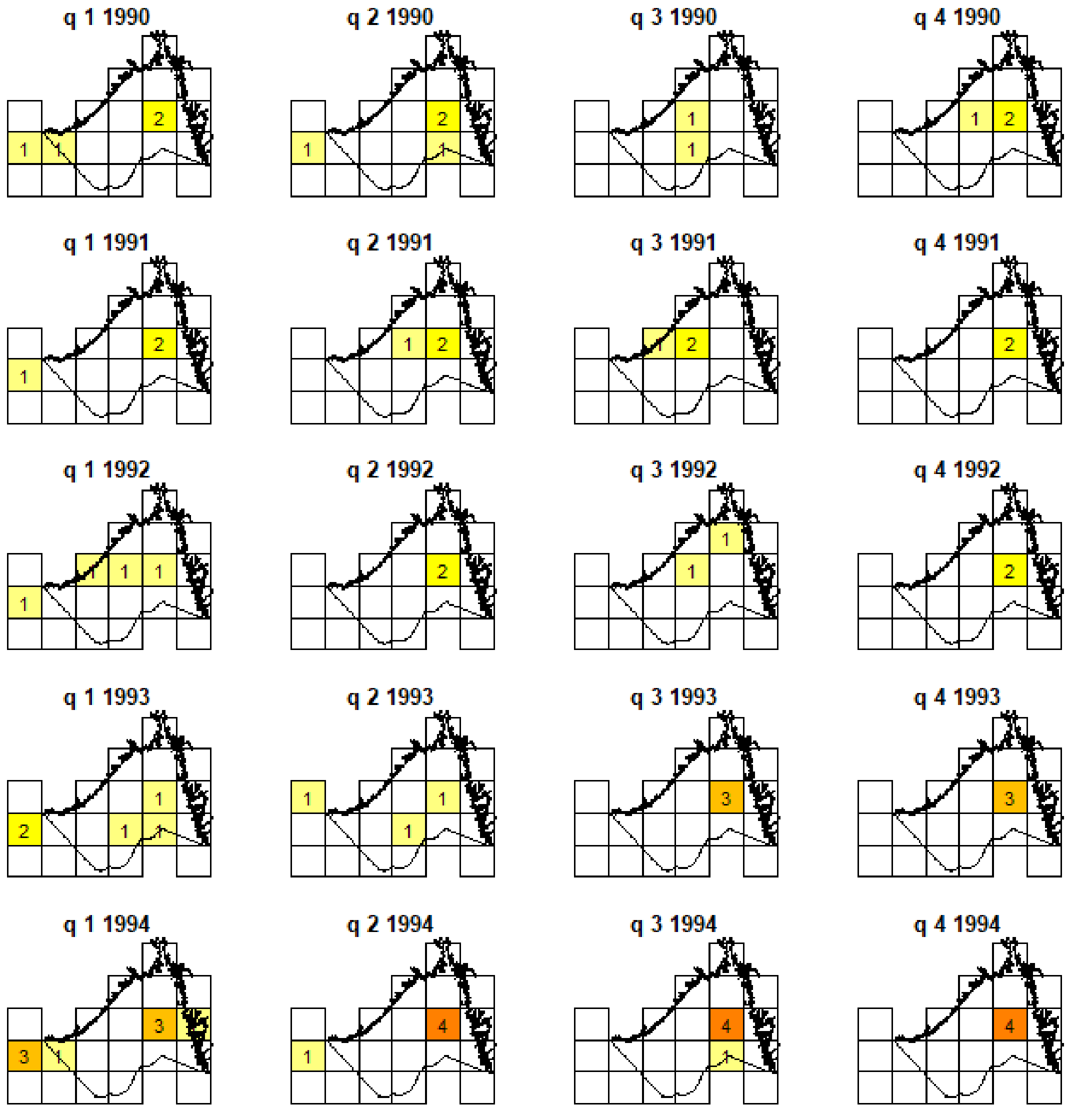
Sampling program	Strategy	Sampled fractions	Start year	End year
22	Observers	sorted (boiled, raw, lus)	2010	2010
	Self-sampling	unsorted	2011	2020
13	Observers*	sorted (boiled, raw)	1990	2004
	Observers*	unsorted	2004	2004
11	Observers	sorted (boiled, raw, lus)	2010	2020

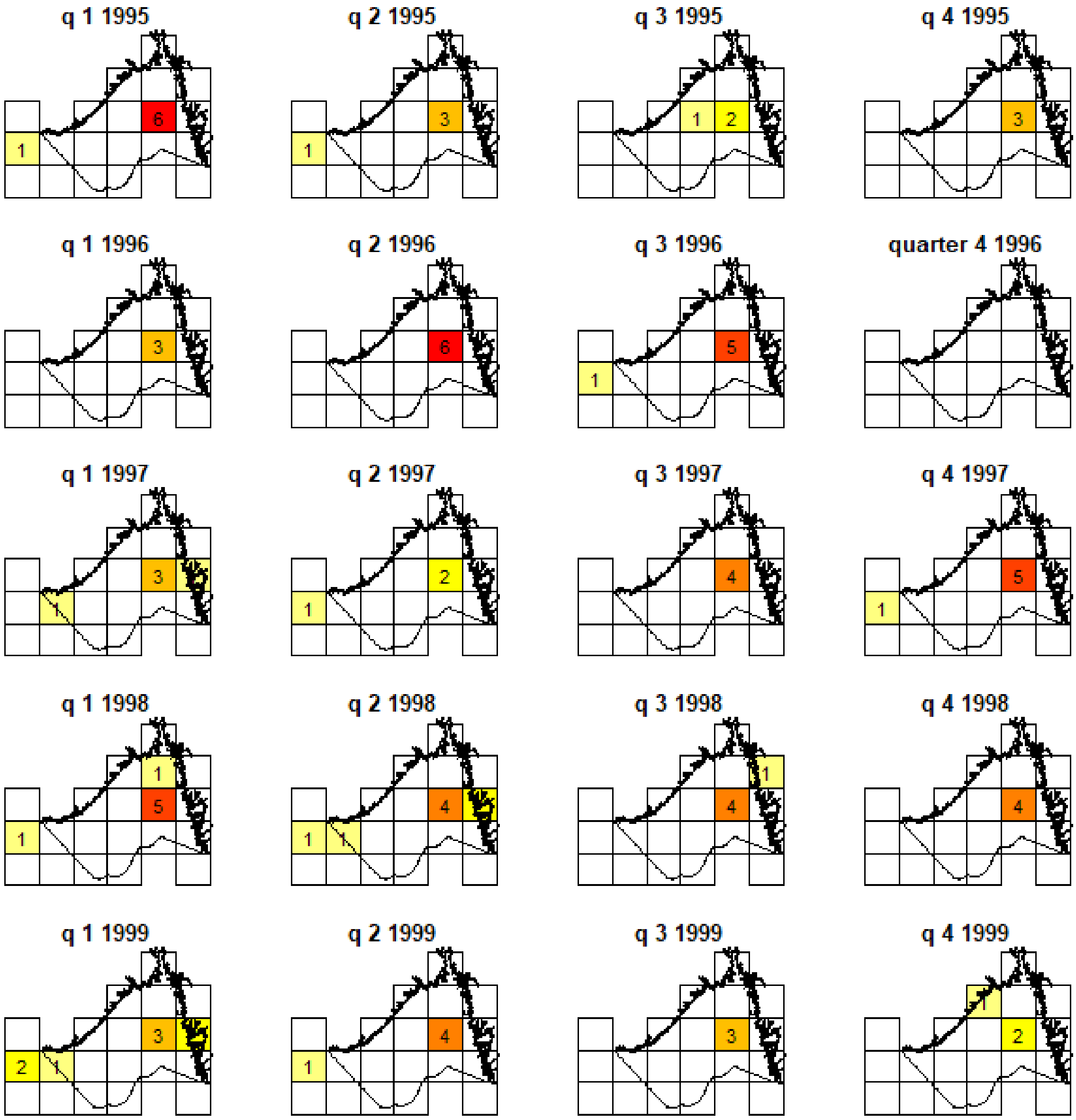
Samples from shrimp catches are for the sampling programs 22 and 11 taken throughout the year by randomly selecting vessels from a list of those being active (described in the methods section above). For the (historical) program 13, at which catches were sampled at the auction in Gothenburg, the method was different in that samples were simply collected from the catch of boats available and willing to assist. The aim of each program was to collect 3 samples, i.e. from three trips, a quarter. For the sorted catch, the sampled fractions (from one haul per trip) were combined together with the unsorted samples into one data set after weighting the sampled fractions according to the total weight of the fraction. Only the samples from Skagerrak were included in the analyses.

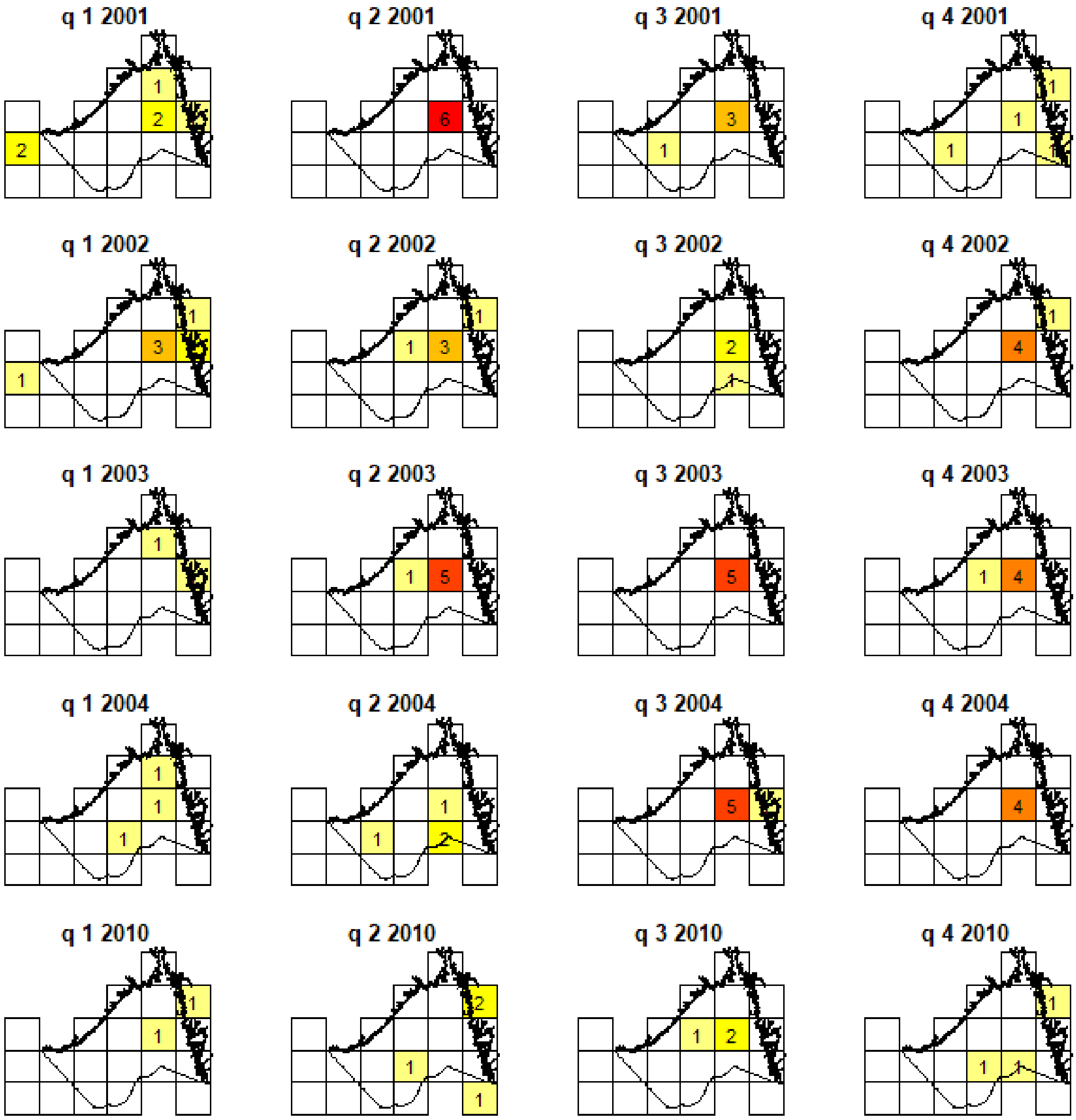
Figures A1 displays the number of hauls/trips sampled per quarter and year for the combined data set. Figure A2 displays the number of hauls/trips per year for all the quarters combined. The figures show clearly the limited and fairly static spatial distribution of the samples from the Swedish commercial shrimp catch.

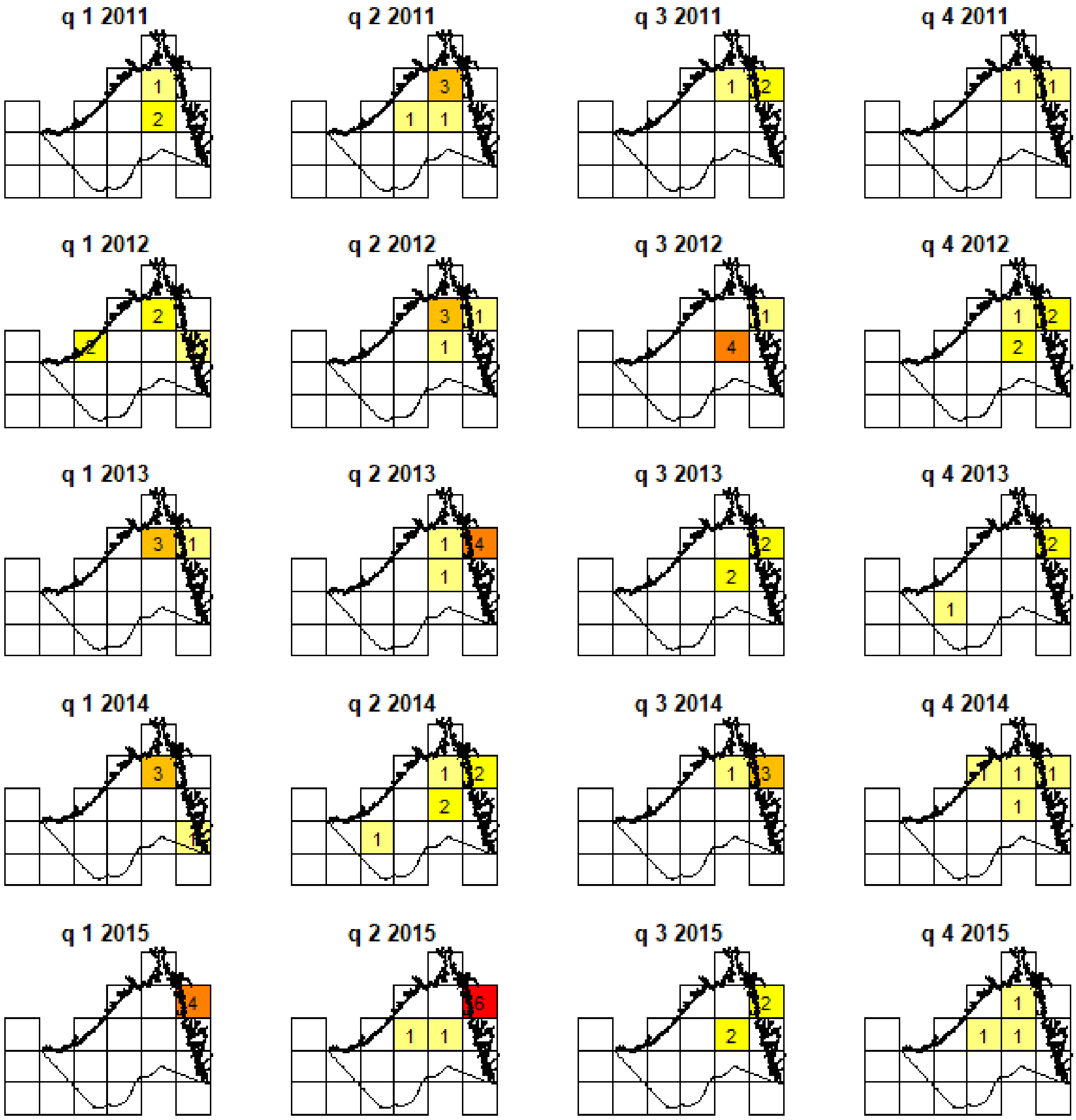
The sampling program 13 only sampled two of the catch fractions (boiled and raw), while the programs 11 and 22 also sampled the lus fraction of the catch, either directly or via the unsorted sample. As an example, figures A3 and A4 illustrates the size distribution of weighted numbers per carapace length per sex and per size fraction respectively. To understand the influence of the lus sample on estimated size at sex change (L50), a scenario was run excluding the lus fraction from all samples, that is, excluding the unsorted catch samples and the lus fraction from the sorted catch, and comparing these to the scenarios including the lus fraction (Figure A5). It is only the time period 2010-2019 which is relevant for the comparison as the only ongoing sampling program before this was program 13, during which lus was not sampled (the exception of year 2004). In the scenario excluding the lus fraction, sample numbers were too few in many years for a reliable estimate of L50. However, for the years of this scenario, but it shows that the lus fraction does not seem to influence the value of L50 markedly, but that the uncertainty around the estimates were smaller when the sample numbers were sufficient (Figure A5).

A1


q 1 1990
q 2 1990
q 3 1990
q 4 1990
q 1 1991
q 2 1991
q 3 1991
q 4 1991
q 1 1992
q 2 1992
q 3 1992
q 4 1992
q 1 1993
q 2 1993
q 3 1993
q 4 1993
q 1 1994
q 2 1994
q 3 1994
q 4 1994


q 1 1995
6
1
q 2 1995
3
1
q 3 1995
1
2
q 4 1995
3
q 1 1996
3
q 2 1996
6
q 3 1996
5
1
quarter 4 1996
q 1 1997
3
1
q 2 1997
2
1
q 3 1997
4
q 4 1997
5
1
q 1 1998
1
5
1
q 2 1998
4
1
1
q 3 1998
1
4
q 4 1998
4
q 1 1999
3
2
1
q 2 1999
4
1
q 3 1999
3
q 4 1999
1
2


q 1 2001
q 2 2001
q 3 2001
q 4 2001
q 1 2002
q 2 2002
q 3 2002
q 4 2002
q 1 2003
q 2 2003
q 3 2003
q 4 2003
q 1 2004
q 2 2004
q 3 2004
q 4 2004
q 1 2010
q 2 2010
q 3 2010
q 4 2010


q 1 2011
q 2 2011
q 3 2011
q 4 2011
q 1 2012
q 2 2012
q 3 2012
q 4 2012
q 1 2013
q 2 2013
q 3 2013
q 4 2013
q 1 2014
q 2 2014
q 3 2014
q 4 2014
q 1 2015
q 2 2015
q 3 2015
q 4 2015

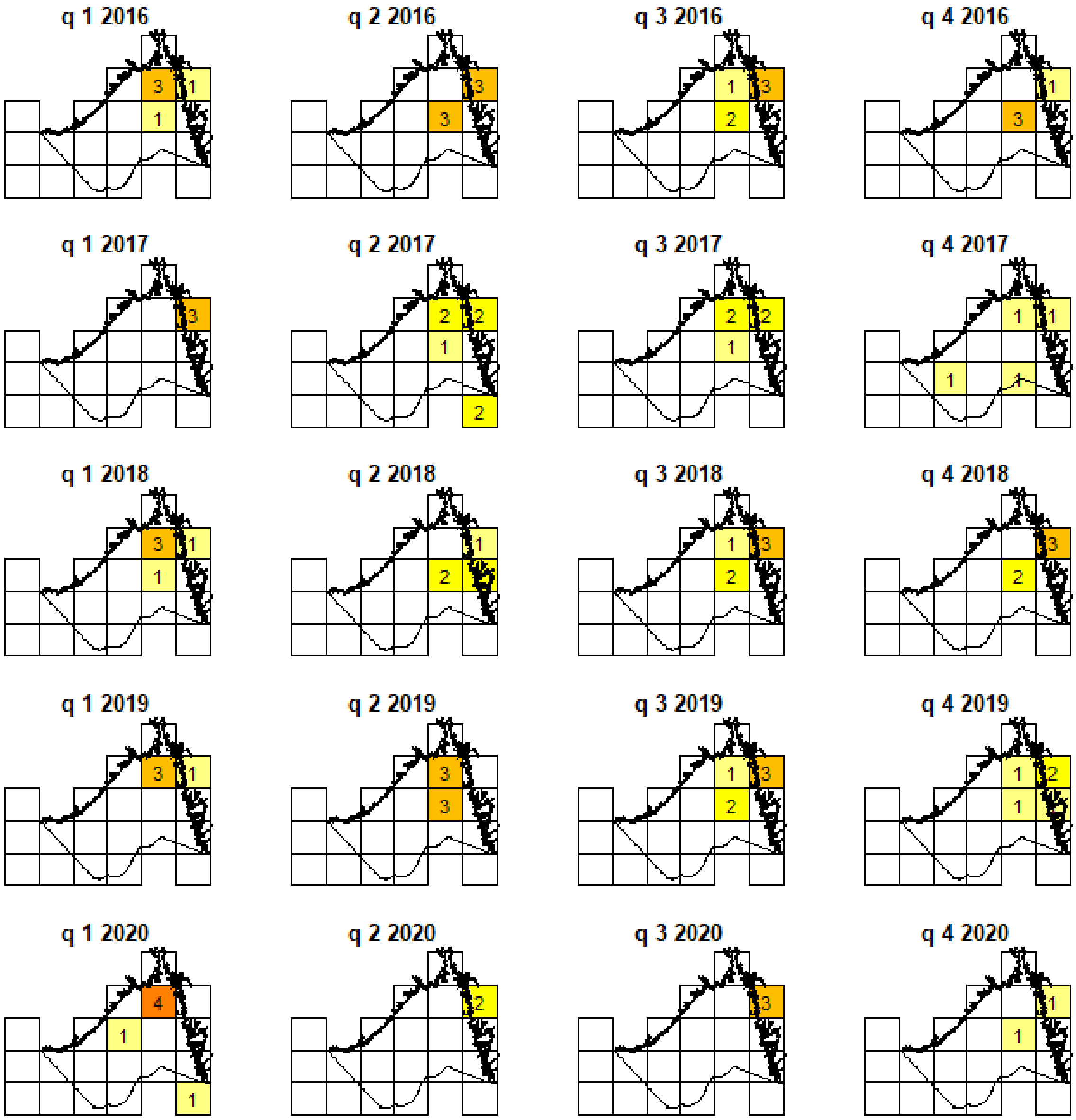
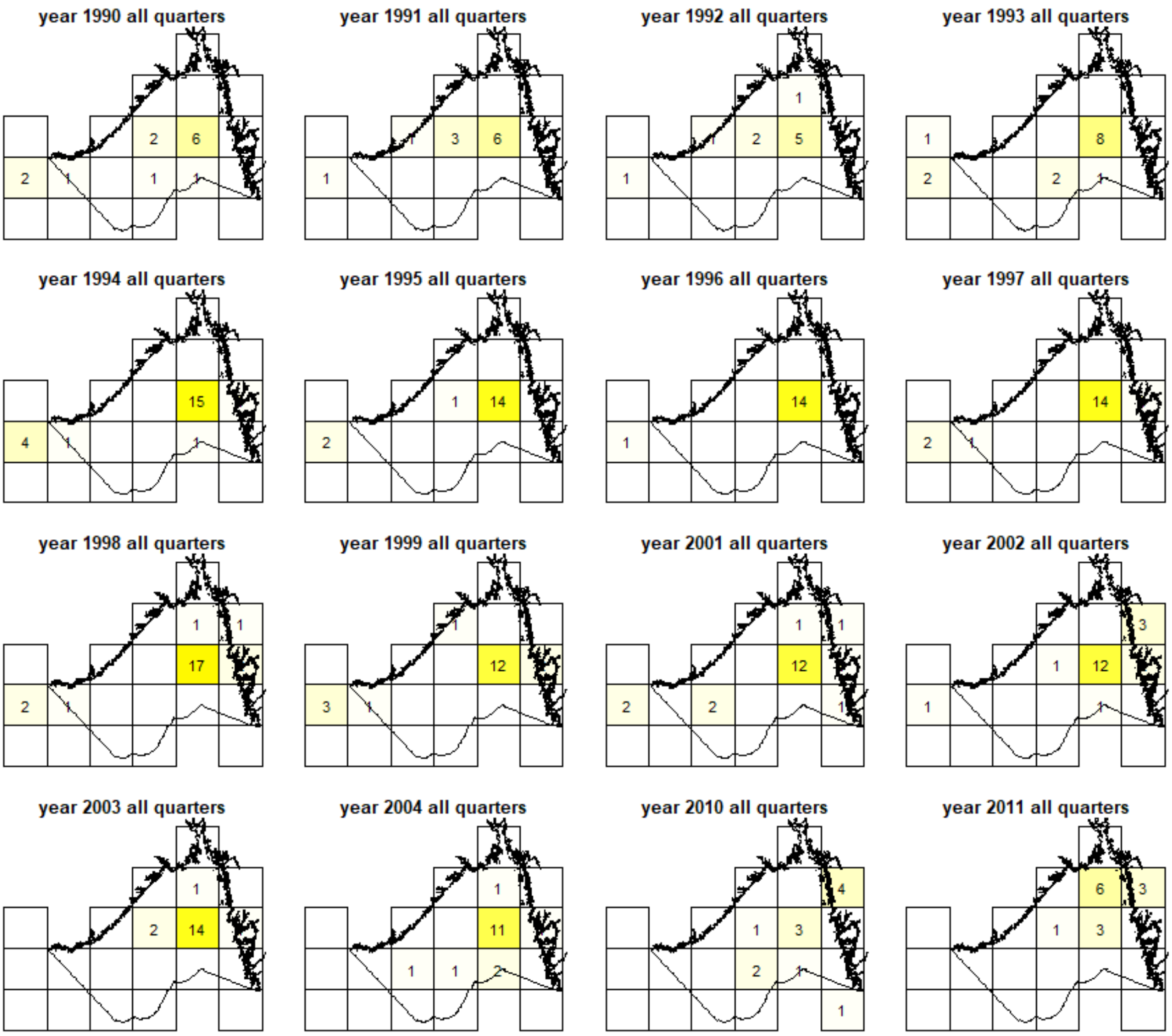


Figure A1. Number of trips/hauls per year per quarter in the data set with samples from all three sampling programs combined.

A2


year 1990 all quarters
year 1991 all quarters
year 1992 all quarters
year 1993 all quarters
year 1994 all quarters
year 1995 all quarters
year 1996 all quarters
year 1997 all quarters
year 1998 all quarters
year 1999 all quarters
year 2001 all quarters
year 2002 all quarters
year 2003 all quarters
year 2004 all quarters
year 2010 all quarters
year 2011 all quarters

A2, contiued

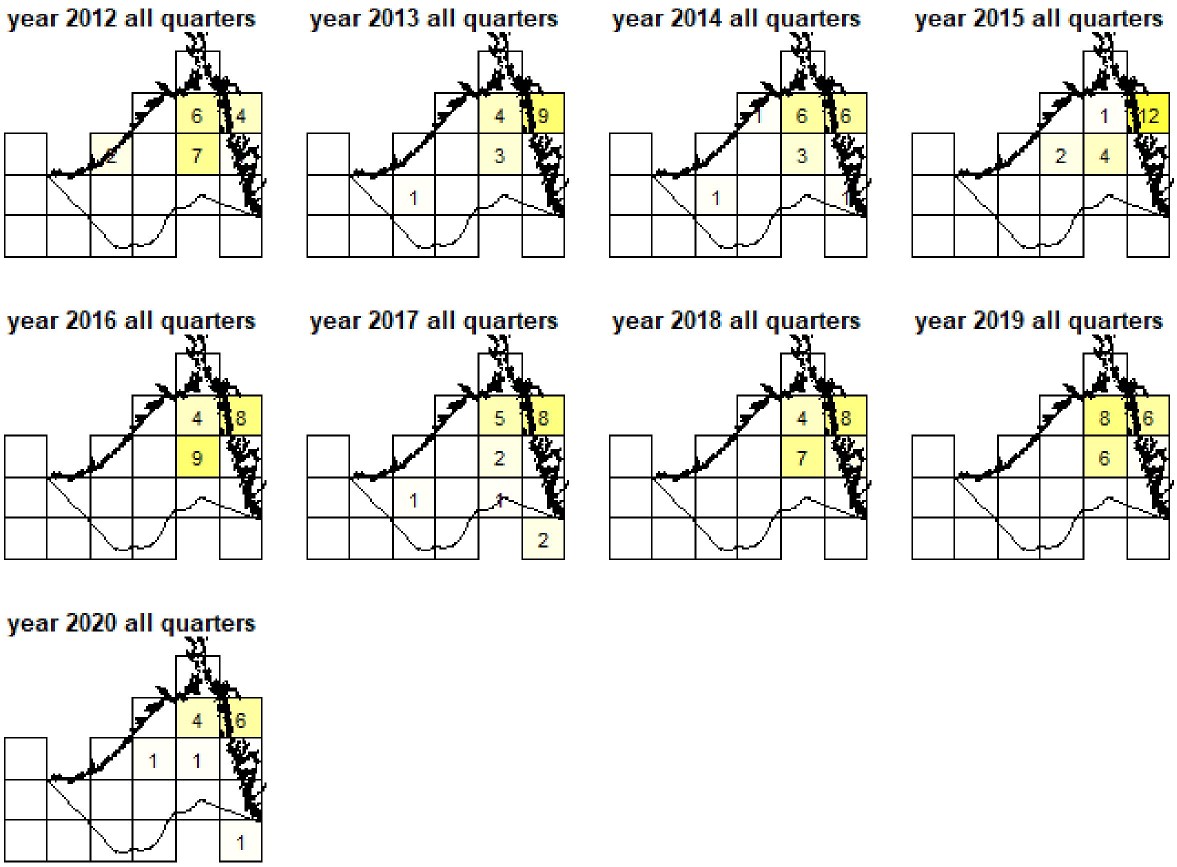


Figure A2. Number of trips/hauls per year in the data set with samples from all three sampling programs combined.

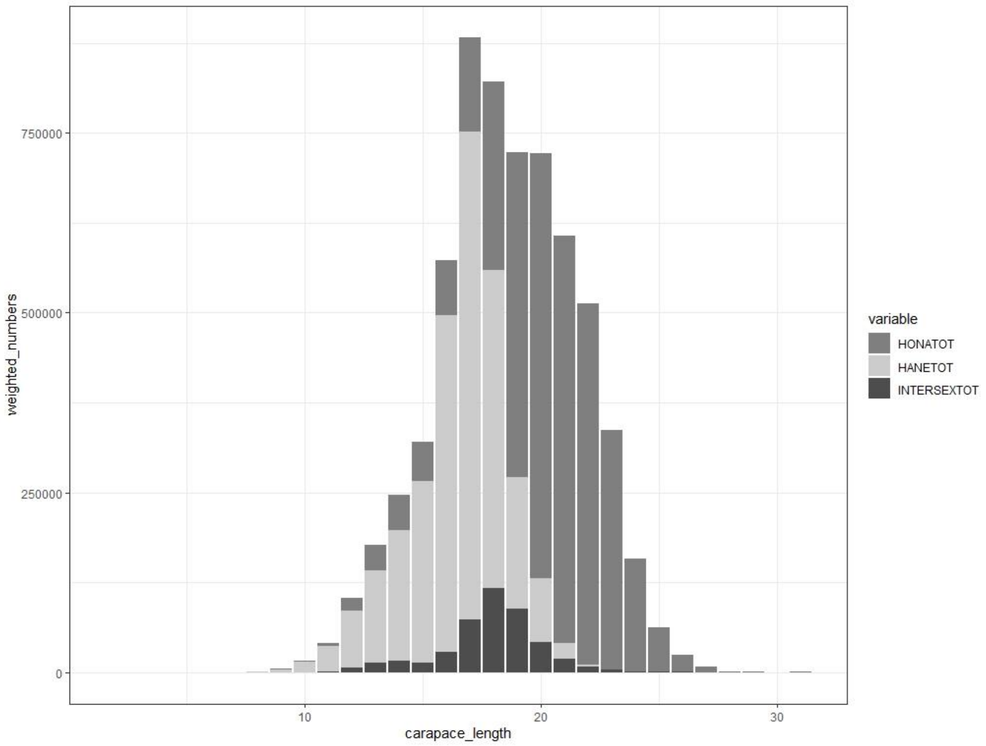


Figure A3. *Pandalus borealis*. Size distribution per carapace length and sex in sampling program 11 for all years combined (2010-2020).

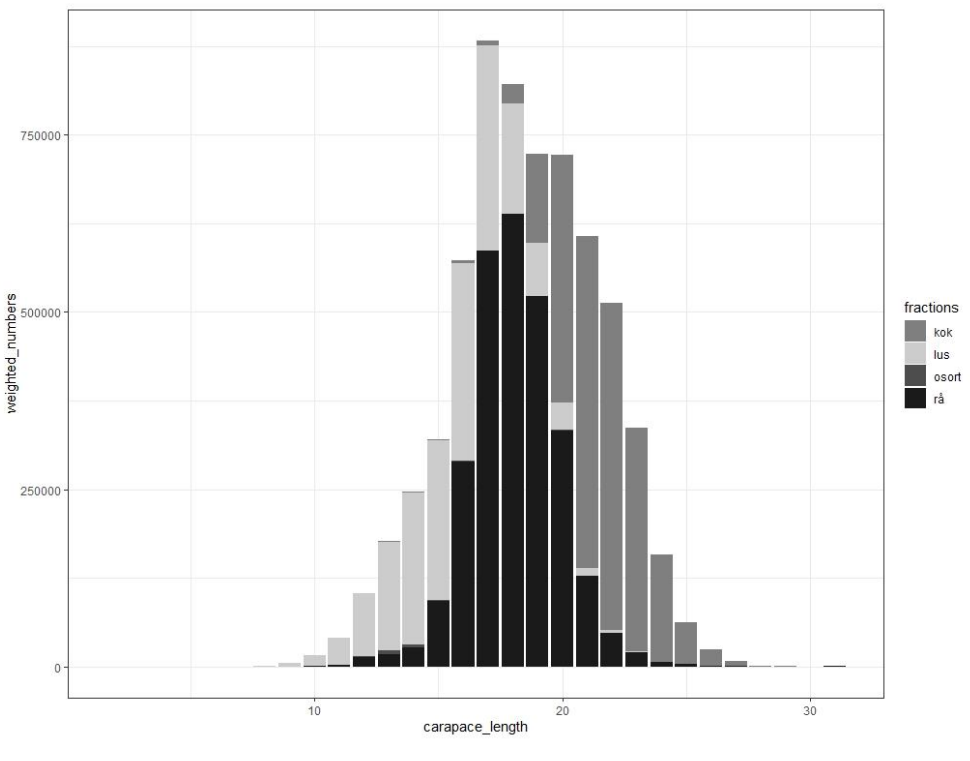


Figure A4. *Pandalus borealis*. Size distribution per carapace length and size fraction in sampling program 11 for all years combined (2010-2020).

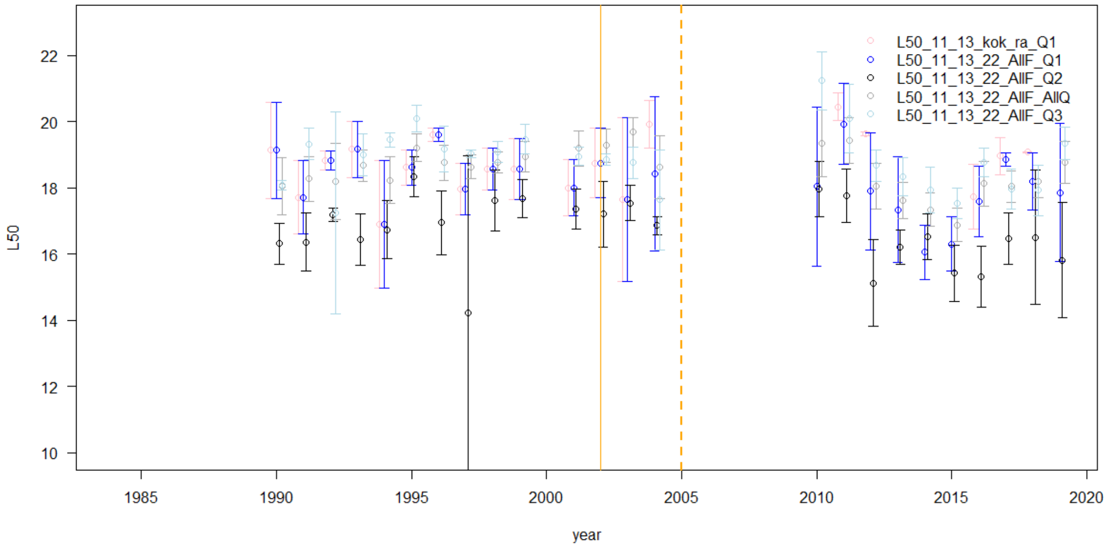


Figure A5. A comparison of annual estimates size at sex change (L50) when excluding (L50_11_13_kok_ra_Q1) and including (L50_11_13_22_AllF_Q1, L50_11_13_22_AllF_Q2, L50_11_13_22_AllF_Q3 and L50_11_13_22_AllF_AllQ,) the lus fraction. The extension_Q1, Q2 and AllQ refers to quarter 1, quarter 2, quarter 3 and all quarters, respectively.

WD10 Estimating a predation index for the 27.3a-4s northern shrimp (Pandalus borealis).

by Patrik Börjesson, Mikaela Bergenius, Guldborg Søvik.

Estimating a predation index for the 27.3a-4s northern shrimp (*Pandalus borealis*).

Patrik Börjesson, Mikaela Bergenius, Guldborg Søvik.

Introduction

In this study, we use a generalized additive model (GAM) to estimate standardized biomass indices for a number demersal species and, in combination with published diet data, to generated combined predation indices for the Norwegian Deep/Skagerrak (NDSK) pandalus stock. An advantage of using modelling instead of design-based methods is that it is easier to include spatiotemporal variation in distribution of the predator species, and to account for differences in catchability between vessel (Maunder and Punt 2004) and/or surveys (Moriarty 2020).

Material and Methods

Survey data

Norwegian shrimp survey

The Norwegian bottom trawl survey for northern shrimp (*Pandalus borealis*) in the Skagerrak and Norwegian deep (ICES Divs. 27.3.a and 27.4.a. east) has been conducted since 1984. Until 2004 the survey was conducted in quarter 4. In 2004 and 2005, the timing of the survey was conducted in May, and in 2006, the survey period was changed to the 1st quarter (January/February) to improve the estimates of spawning stock biomass (SSB) and recruitment of northern shrimp. In addition to data on shrimp lengths and sex, data on multiple demersal fish species preying on the northern shrimp are collected. The survey is based on a stratified fixed design, covering depths from 100 – 550 meters in four depths zones (100-200m, 200-300m, 300-500m and >500m, see Figure 1). The stratification has been revised over time and the number of strata has been reduced but the depth zones have been maintained. The list of sampling stations has also been revised and currently consists of 111 fixed stations, of which 60 stations are located in the Skagerrak.

The Norwegian shrimp survey uses a 'Campelen 1800' trawl with 14-inch rockhopper. Up until 2017 the mesh size of the codend was 20 mm, since then 10 mm mesh size is used. From 2008 and onwards strapping has been used to keep gear geometry constant, with an expected doorspread of 48-52 meters regardless of depth. With a trawl duration of 30 minutes and a speed of 3 knots, the swept door area will be approximately 0.14 km2.

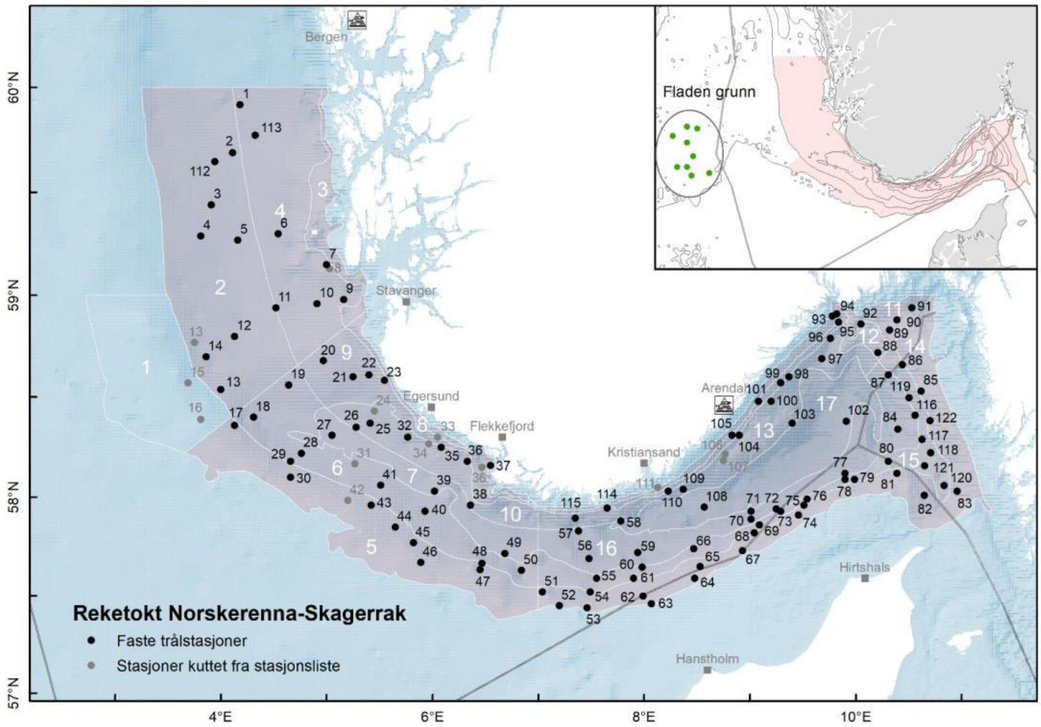


Figure 1. Norwegian shrimp survey in Skagerrak and the Norwegian Deep (ICES Divs. 3.a and 4.a East): Strata system with the 111 fixed trawl stations. Trawl stations marked in grey have been deleted from the station list. Trawl stations on Fladen Ground were fished 1978-1994 and 2021. *(Map from Søvik and Thangstad 2021).*

The International Bottom Trawl Survey (NS-IBTS)

The International bottom trawl survey in the North Sea (NS-IBTS) has been conducted in the 1st quarter of the year since the early 1960s. Initially the survey was focusing on juvenile herring, but was later broadened to include recruitment of gadoids. The Skagerrak/Kattegat was included in the 1980s and the quarter 3 survey was initiated in 1991. The main objectives for the survey are to collect data for recruitment indices of the main commercial species and to monitor distribution and relative abundance of all fish species and selected invertebrates (ICES 2020).

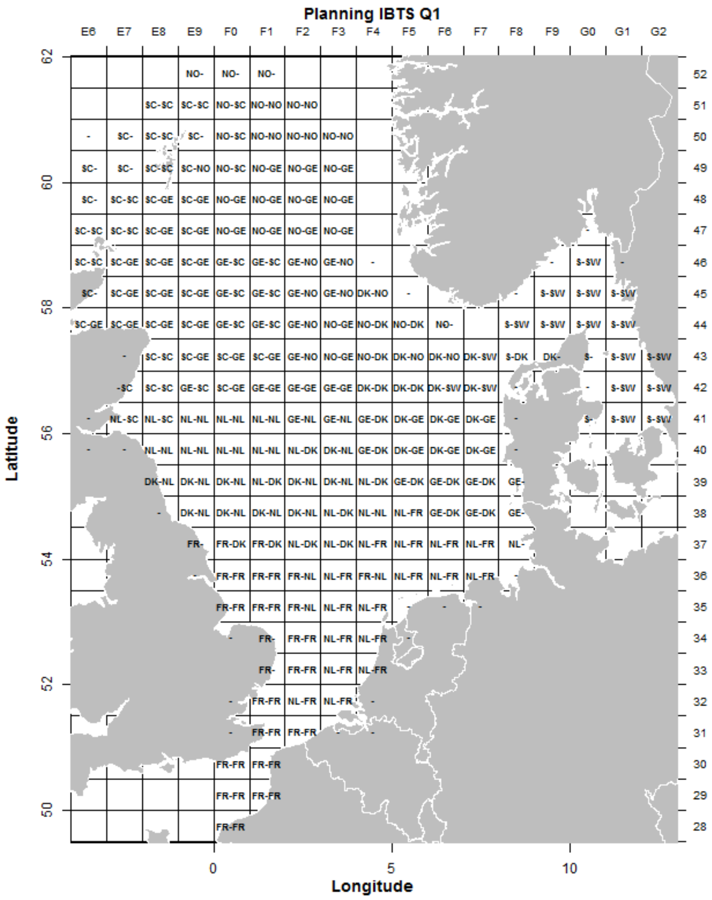


Figure 2. Survey area for the IBTS in quarter 1. Map showing survey grid and present allocation of rectangles between countries (ICES 2020)

The NS-IBTS is stratified based on ICES statistical rectangles (1 degree longitude x 0.5 degree latitude, Figure 2) covering depths from 20-250 meters. Each rectangle is typically sampled once by two different countries so that two hauls are taken per rectangle. Stations are selected from a common haul data base or from national list of previously identified 'clear' trawls tracks. Random selection is supported, but a main criteria is that hauls stations should not be clustered or to close in space or time. Skagerrak and Kattegat is mainly fished by Sweden and a list of fixed station are used, currently 27 hauls in Skagerrak and 19 hauls in Kattegat. The Swedish quarter 3 survey was following the same protocol until 2005 when the design was changed to a semi-random depth stratified design with seven depth strata (21-40m, 41-60m, 61-80m, 81-100m, 100-150m, 150-200m & 200-300m). The number of station allocated to each depth strata is proportional to the size of the strata and randomly selected from a national haul database.

The NS-IBTS uses the GOV trawl (chalut à Grande Ouverture Verticale) which is a larger gear than the Campelen trawl. No strapping is used but different sweep lengths (60/110 m) have previously been used for shallow (< 70 m depth) and deep hauls in quarter 1 to compensate for differences in warp length at different depths. In recent years only Sweden and Norway uses the long sweeps for deeper hauls in quarter 1. Short sweeps have always been the rule for the NS-IBTS Q3 survey. Door spread varies from approximately 70 – 110 meter. Given the nominal trawl duration of 30 minutes at 4.0 knots the expected swept areas range from 0.26 -0.41 km2.

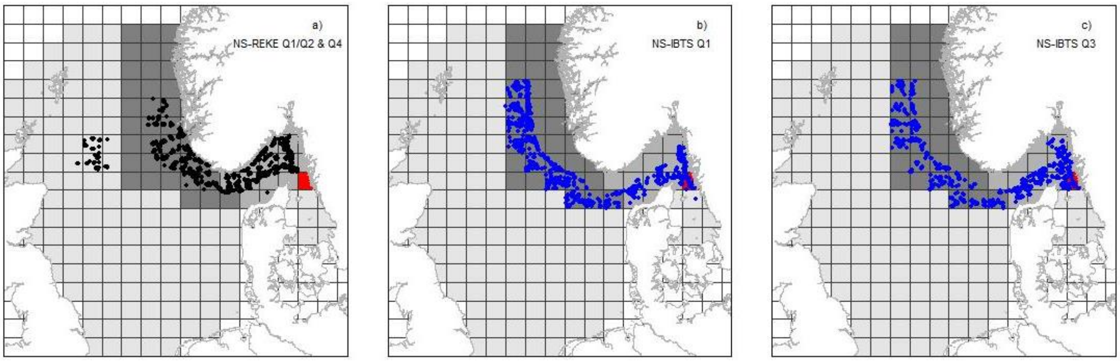


Figure 3. Distribution of survey data used in the analyses. a) The Norwegian shrimp survey 1984-2021, b) the NS-IBTS Q1 1984-2021, and c) the NS-IBTS 1991-2021. Dark and medium dark shading represents the Norwegian Deep and the Skagerrak parts of the NDSK pandalus distribution, respectively. Light grey shows the survey area of NS-IBTS. The red square was included as a part of Skagerrak in the NS-IBTS survey. The hauls from Fladen Ground is not used in the analyses.

The Norwegian shrimp survey and the NS-IBTS overlap along the 100-250 meter deep contours that run through the central parts of the Skagerrak and the Norwegian depths. The two surveys are otherwise spatially separated, where the Norwegian shrimp survey covers the deeper parts of the Skagerrak and the Norwegian depths, and NS-IBTS extends to the south and to the west.

Access to data and data preparation

Raw data from the Norwegian shrimp survey 1984-2021 were provided by the Institute of Marine Research in Norway (Søvik 2021). In addition, we used biomass indices, calculated as average catch in kg per towed nm, for 21 species from 1984 to 2021. The list also contains a combined estimate for skates and rays and, from 2006, the blackmouth catshark (*Galeus melastomus*) (ICES 2005, 2006, Søvik and Thangstad, 2021).

NS-IBTS data for the 1st quarter 1984-2020 and the 3rd quarter 1991-2020 were downloaded from DATRAS (https://datras.ices.dk/Data_products/Download/Download_Data_public.aspx) using the exchange format for biological data (HL-table) and the swept area assessment product prepared by WKSAE_DATRAS/WKABSENS (2021a,b) for haul information (HH-table). The data product span from 1967 to 2020 so for 2021 the data product was supplemented with haul information from DATRAS.

Valid haul information and catch data were filtered out from the data sets based on statistical rectangles corresponding the distribution of the NDSK pandalus stock (Figure 3) and converted to common units to generate three data tables, containing haul information, catch information and numbers at length. The tables were linked with a unique haul ID carrying information on year, survey, vessel and station. Auxiliary data such as swept area, doorspread and trawled distance were screened for outliers and missing values.

Pandalus predators

In 2014, Jørgensen et al compiled available information on shrimp in the stomach contents of fish species in Skagerrak and the Norwegian Deep. The data came from a series of 13 survey cruises carried out in the area 1984 – 1987 (Jørgensen et al 2014 and references therein). Based on the fraction pandalus, measures as % wet weight, they identified four of the 21 species reported by the Norwegian shrimp survey as primary pandalus predators; blue whiting (*Micromesistius* *poutassou*), roundnose grenadier (*Coryphaenoides rupestris*), velvet belly (*Etmopterus spinax*), and cod (*Gadus morhua*). The remaining species was classified as incidental predators or as having unknown prey preferences, but in their final analyses the only used the four primary predators. Available data from 1984 to 1991 in ICES stomach database (ICES 2010) support the classification of cod as a pandalus predator over a wider area during the period, but also indicate that saithe (*Pollachius virens*) may be of equal importance.

Completely comparable data from later years are not available, but Skorda (2018) reported that pandalus occurred most frequently in stomach contents of roundnose grenadier, cod, saithe and blue whiting in a sample from the Norwegian shrimp survey 2017, but also that is was common in the stomachs of greater argentine (*Argentina silus*), anglerfish (Lopius *piscatorius*) , poorcod (*Trisopterus esmarkii*) , redfish (*Sebastes* spp), whiting (*Merlangius merlangus*) and haddock (*Melanogrammus aeglefinus*). Unpublished data from the Norwegian shrimp survey in 2020 also lend support to the classification of roundnose grenadier, velvet belly, blue whiting and saithe as pandalus predators. The unpublished data did not support the observation of Skorda (2018) that Greater argentines frequently feed on pandalus. Cod was not sampled for stomach contents during the 2020 cruise.

To illustrate the differences in choice of predators we modelled biomass and predation indices for two assemblages of pandalus predators. We used a four species model based on the primary predators identified by Jorgensen at al 2014, i.e., the three deep-sea species roundnose grenadier, velvet belly and blue whiting, plus the Atlantic cod. We also used a nine species model which added starry ray (*Amblyraja radiata*), hake (*Merluccius merluccius*) and three gadoid species; haddock, whiting and saithe. The added species were chosen because: 1) on average they made up a significant fraction of the total catch in the Norwegian shrimp survey, 2) they were identified by one or several sources as pandalus predators, and 3) it was possible to estimate their relative importance as predators (in %W) from the sources used by Jorgensen et al (2014). The species and their relative importance as pandalus predators (in %W) is presented in Table 1.

Only length classes >20 cm were included in the analyses. Mean weight at length was estimated using area-specific weight-length regression based on NS-IBTS data from 1991-2021 and used to calculate catch by species for the truncated length distribution. For blue whiting and starry ray we only included specimens larger than 30cm and 40 cm, respectively (Table 2).

Table 1. Proportion of *Pandalus borealis* (in % wet weight) and frequency of occurrence (%FO) of predator species used for the index. Modified after Jorgensen et al (2014).

Species	% W	% FO	N fish	N Stn	Source
Amblyraja radiata > 40 cm	*8.1*	*7.3*	*58*	*41*	*Skjaeraasen and Bergstad 2000*
Coryphaenoides rupestris	*11.2*	*	*373*		*Bergstad et al 2003*
Etmopterus spinax	*11.9*	*	*84*		*Bergstad et al 2003*
Gadus morhua	*23.6*	*44.8*	*777*	*378*	*Bergstad 1991*
Melanogrammus aeglefinus	*2.0*	*1.4*	*638*	*122*	*Albert 1994*
Merlangius merlangus	*1.4*	*2.1*	*138*	*36*	*Bergstad 1991*
Merluccius merluccius	*2.7*	*20.4*	*47*	*23*	*Bergstad 1991*
Micromesistius poutassou >30 cm	*9.8*	*7.5*	*215*	*52*	*Bergstad 1991*

Pollachius virens	*2.6*	*8.2*	*1307*	*137*	*Bergstad 1991*

*) Frequency of occurrence not presented in source.

Tale 2. Weight length parameters used for estimating weight of predator species. All parameters were estimated from fish caught in the Skagerrak, ices area 27.3.a.20.

Species	log a	b	N fish	Range (cm)	Source
Amblyraja radiata	*-11.76*	*2.99*	*188*	*5-121*	*IBTS 1991-2021*
Coryphaenoides rupestris	*-8.11*	*2.91*	*234*	*4-21*	*SE Skagerrak survey 2018-2021*
Etmopterus spinax	*-12.11*	*2.90*	*77*	*11-54*	*SE Skagerrak survey 2018-2021*
Gadus morhua	*-12.08*	*3.14*	*11513*	*5-121*	*IBTS 1991-2021*
Melanogrammus aeglefinus	*-12.14*	*3.17*	*5491*	*6-70*	*IBTS 1991-2021*
Merlangius merlangus	*-12.06*	*3.09*	*4234*	*4-56*	*IBTS 1991-2021*
Merluccius merluccius	*-12.26*	*3.08*			*IBTS 1991-2021*
Micromesistius poutassou	*-12.66*	*3.20*			*IBTS 1991-2021*
Pollachius virens	*-11.99*	*3.08*	*2483*	*10-110*	*IBTS 1991-2021*

Index calculation

We used a delta-gam approach to standardize the catch per unit effort (*cpue*) of 9 fish species identified as important or potentially important predators on pandalus with similar specification for 1) the presence/absence (binomial) model, and 2) for the positive (lognormal) model:

1) presence/absence ~ year + survey + te(lon, lat, quarter) + s(depth), family=binomial
2) log(kg) ~ year + survey + te(lon, lat, quarter) + s(depth), offset= log(swept_area_door_km2), family=Gaussian(link='identity)

Year and survey was treated as fixed factors in the model. The spatiotemporal component, te(lon, lat, quarter), was modelled as a two-dimensional thin plate regression spline, including quarter as an 'interaction term'. Depth was modelled as a one-dimensional thin plate spline and the log of swept area in square kilometres was include as an offset to account for variation in effort between fishing events. Survey (NO-SHRIMP & NS-IBTS) were included as a factor in the model to account for survey-gear effects, assuming that the difference between trawls and ground gears (Campelen vs GOV) would be more important than differences between ships within the NS-IBTS. The Norwegian survey fish during both day and night time, whereas IBTS only fish during daytime. This should probably also be included in later analyses.

A prediction grid for the stock area with the resolution 5/8 × 5/8 arc minutes (ca 1160 meter grid) was created based on spatial depth information from EMODNET (2018), and values was predicted for the depth range 80 – 500 m in quarter 1. Predicted values were back-transformed and used to calculate the standardized catch per unit effort in kg/km^2 by year and area (the North

Sea and Skagerrak parts of the survey area respectively). The modelling were using the R-package 'mgcv' (Woods 2011).

From the standardized single species *CPUE*s we calculated two indices, *i*) a combined biomass index (*BI)*, and *ii*) a combined predation index (*PI*). The combined biomass index is simply the sum of the standardized single species *CPUE*s, by year and area, whereas the predation index also takes the importance of different pandalus predators from Table 3 into account.

Results and discussion

From 1984 – 2021, 7071 valid hauls in northeastern North Sea and the Skagerrak were available for analyses. 50% of the available hauls came from first quarter, and about 25% from quarter 3 and quarter 4 respectively. In the Norwegian shrimp survey the number of valid hauls have varied between 43 hauls in 2006 to 113 in 2019 when all stations were sampled (Table 3). Data from the 2003 & 2016 survey were invalidated due to technical problems during the survey. In the IBTS the number of hauls have been relatively stable throughout the time period with 40 – 70 stations sampled each in each cruise with the exception of year 2000 when the Swedish IBTSQ3 was cancelled (Table 3).

Table 1 Number of hauls included in the analyses by year, survey and quarter from 1984 to 2021.

	NO-SHRIMP			NS-IBTS	
	Q1	Q2	Q4	Q1	Q3
1984			62	60	
1985			82	54	
1986			62	63	
1987			104	70	
1988			114	53	
1989			97	52	
1990			77	56	
1991			103	67	46
1992			99	55	45
1993			108	54	48
1994			110	58	47
1995			102	56	45
1996			103	54	43
1997			93	60	43
1998			95	59	45
1999			97	58	58
2000			98	58	29
2001			69	57	60
2002			77	58	61
2003			*)	56	57
2004		60		56	57
2005		84		61	61
2006	43			53	51
2007	64			54	56
2008	73			58	57
2009	91			59	46
2010	95			54	54
2011	89			56	60
2012	63			57	57
2013	101			54	61
2014	69			41	61
2015	89			55	65
2016	*)			54	71
2017	108			56	60
2018	110			56	59
2019	113			56	59
2020	104			56	62
2021	121			64	60
Total No Hauls	1333	144	1752	2158	1684

*) Data not available for analyses

The model component describing the positive catch explained from about 40 to 70 % of the deviance with the exception of cod where the model only explained 28.3% of the deviance. Visual inspection of the residuals and fitted vs observed values suggested that most of the models adequately described the data (see example in figure 4), although starry ray showed a quite skewed distribution (Annex 1a).

The effect of position and depth was highly significant whereas the importance of year varied between species (see cod example in figure 5). For all but two species there was also a significant survey effect, with higher catch per unit effort in the Norwegian shrimp survey. The two exceptions were roundnose grenadier and cod. That there would be no effect of survey on the catch of roundnose grenadier seems counterintuitive since the depth range where the species occur, 300 – 600 m (Bergstad 1990) is only sampled by the Norwegian survey. But the Skagerrak stock is at very low levels which is probably the reason no effect of survey was detected.

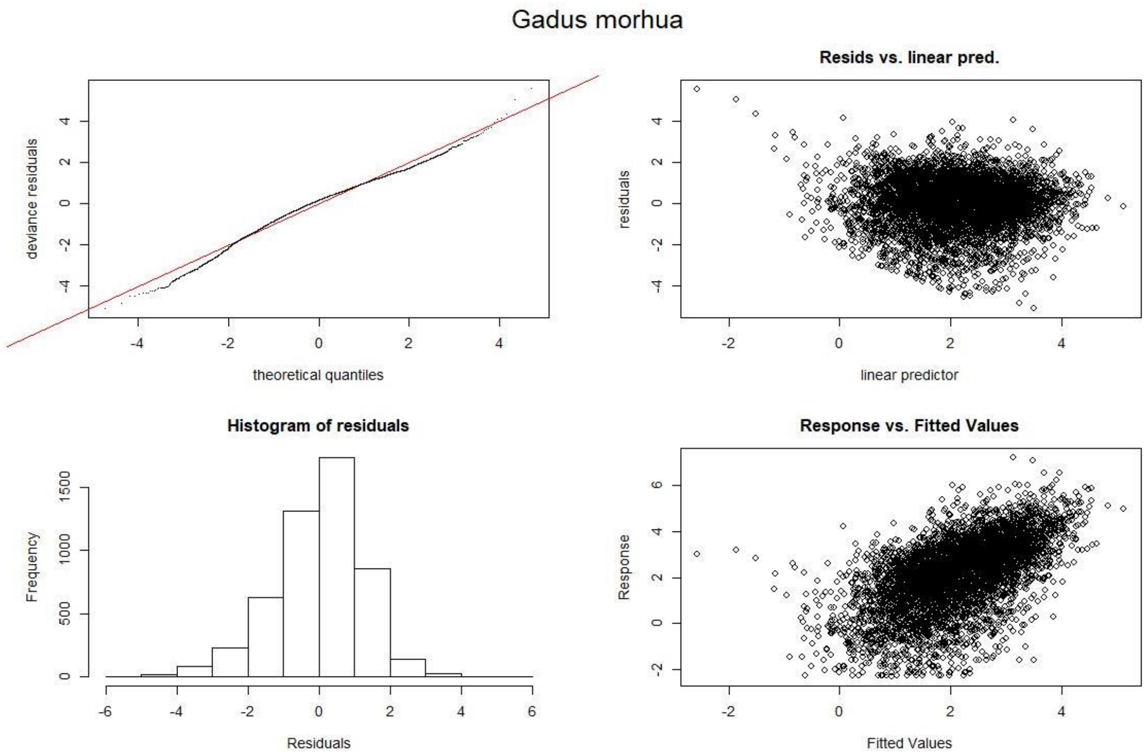


Figure 4. Diagnostics for the delta gam model using cod, Gadus morhua as example. Panel's to the left show distribution of the residuals, top panel to the right show homogeneity of variance and bottom panel to the right visualize the fitted vs observed values which in a perfect fit would be a straight line.

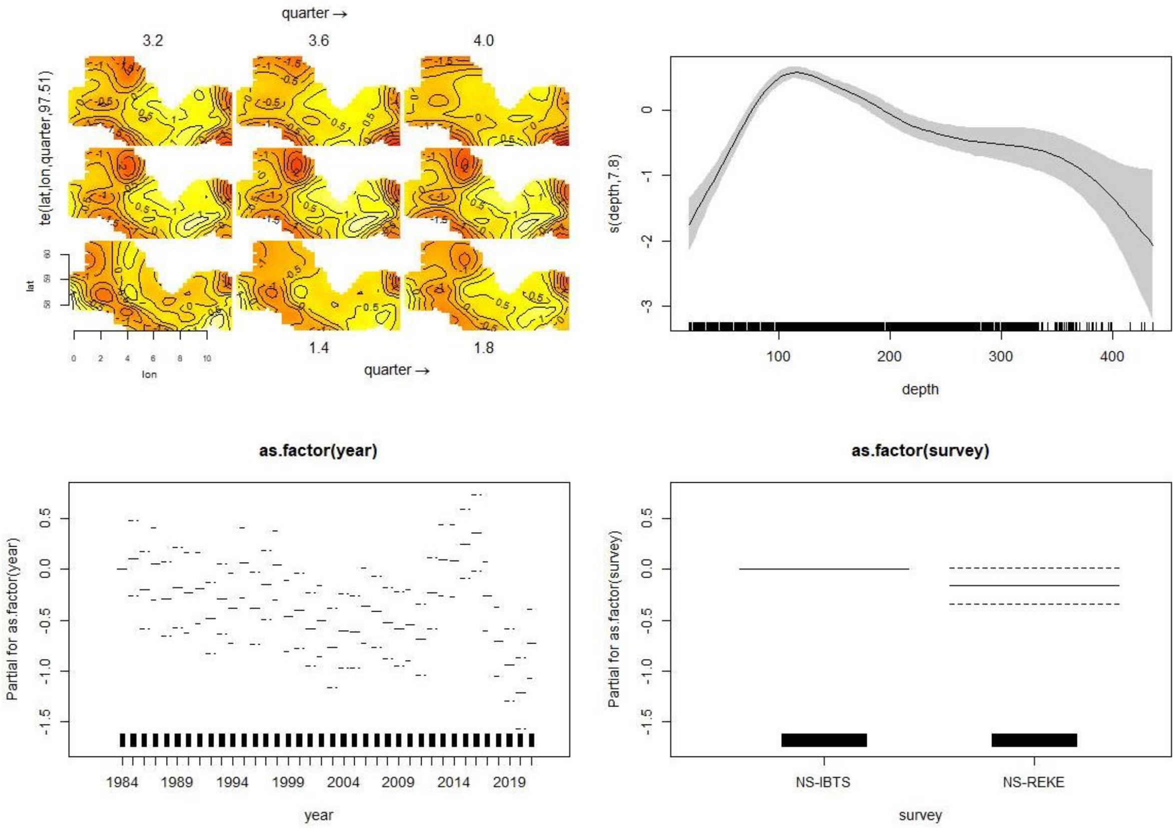


Figure 5 Main effects for the delta gam model using cod, Gadus morhua as example.

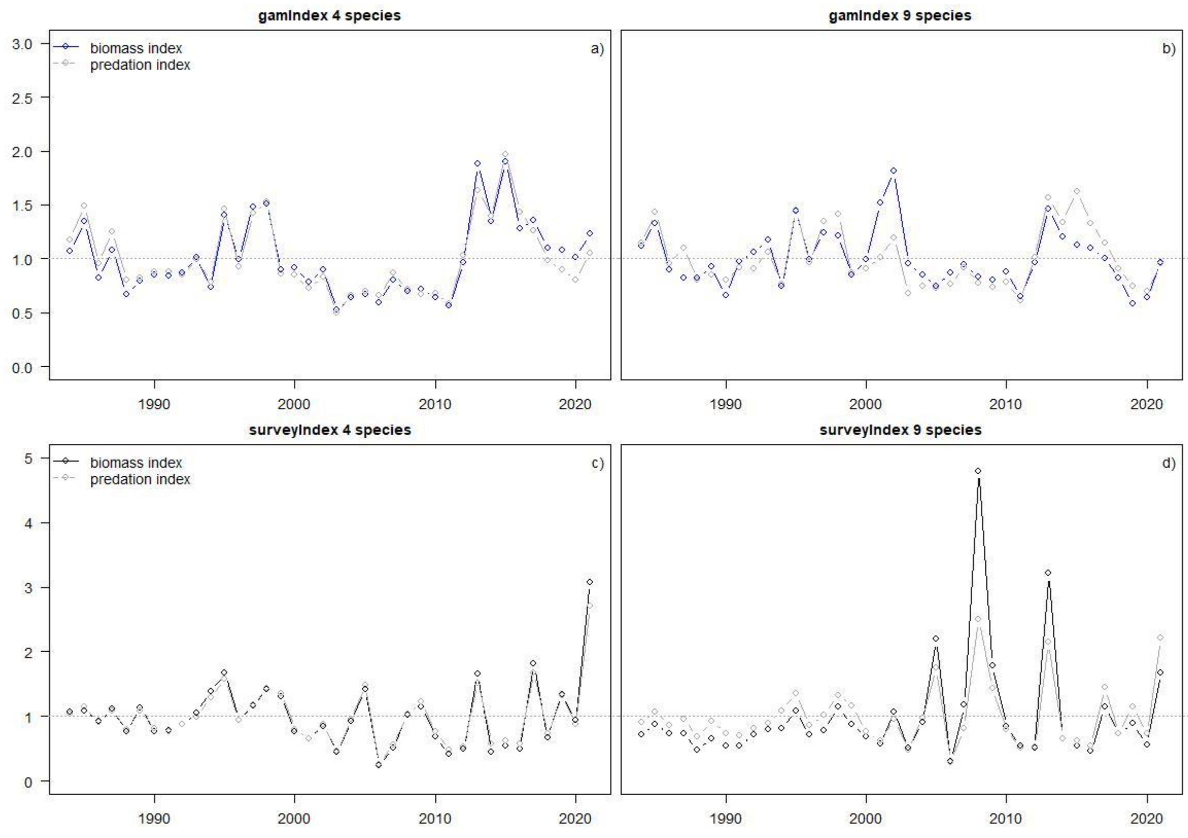


Figure 6. Comparison of the biomass- and predation indices from 1984-2021 based on a) gam model using 4 predators, b) gam model using 9 predators, c) survey index using 4 predators, and d) survey index using

9 predators. Survey data for 2003 and 2016 were not available for modelling but were present in the survey index. Note the scale difference between the upper and lower panel.

Comparison of biomass and predation indices show that the choice of predator composition has an appreciable impact on the index, but that the weighting of predator to get to the predation index contributes little to the final variation (Figure 6). A caveat here is that we used fixed weights for the entire time series. Prey preferences may vary depending of the size of the pandalus stock, the availability of other, potentially preferred prey and the predator's size distribution. Unfortunately time series of prey preferences of predators in the Skagerrak and Norwegian deep are not available.

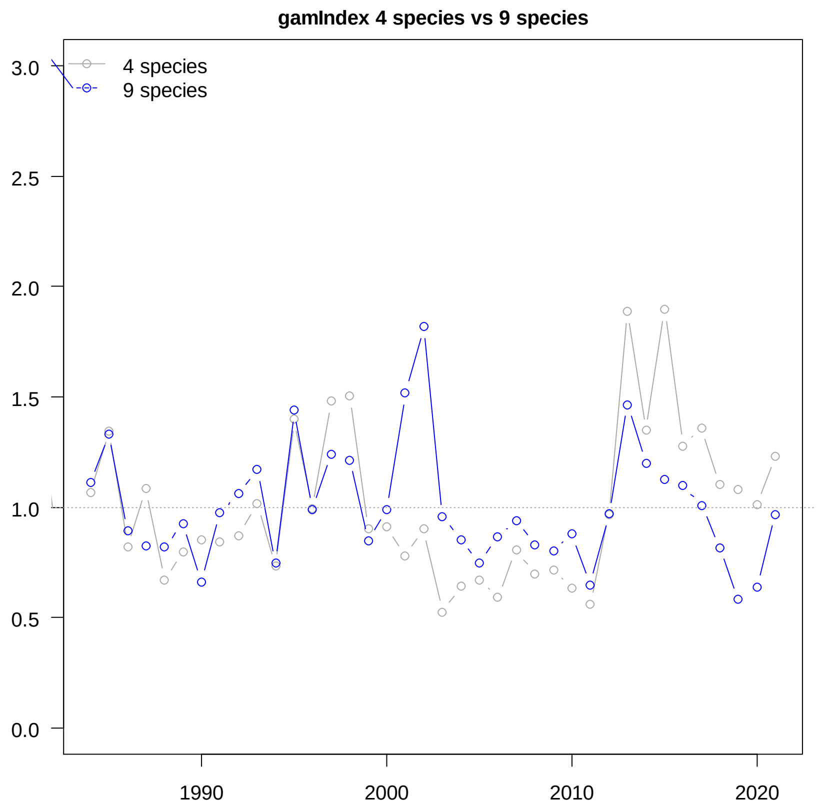


Figure 7. Comparison of the modelled biomass indices 1984-2021 based on a) gam model using 4 predators, b) gam model using 9 predators. Survey data for 2003 and 2016 were not available for modelling but were present in the survey index.

In general, the four species index and the nine species index showed similar trends, but with some notable differences. In 2001 and 2002 the index of the nine species model reached its highest values in the time series while at the same time the four species index was decreasing indicating a lower than average predator biomass. From 2011 to 2013 both indices showed a steep increase, but while the nine species index started to decline immediately after, the four species index peaked again in 2015. The difference between the models probably relates to the species composition. The three deep sea species used in the models is mainly caught at depths below 250 meters which is the depth limit for the NS-IBTS. In practice this means that the four species index is a modelled shrimp survey index, with added data on Atlantic cod from the NS-IBTS. In 2001 and 2002 the cpue index for haddock and whiting peaked (Annex 1e, f) and the cpue of saithe was also going up (Annex 1 i). None of these species are included in the 4 species index. In 2015 and

2016, the cod biomass reaches its highest values in the time series which is reflected in the 4 species index where cod is the only gadoid and roundnose grenadier at a historic low (Annex 1b, d). But in the 9-species index the peak in cod biomass is hidden by other gadoids, contributing to the combined index. This is actually on of the few occasions where the predation index clearly differs from the biomass index in Figure 6a.

The modelling exercise indicates that it is worth combining data from the various surveys to generate a single biomass or predation index for the Norwegian Deep / Skagerrak pandalus population (NDSK). But further work is required before the approach can be implemented in assessment. First, an evaluation of uncertainties in the model is needed. Second, model simplification and data pruning should be evaluated to improve model fit and prediction. Finally, the composition of predators going into the index needs to be settled.

Summary

In this study we modelled biomass- and predation indices for two assemblages of predators on pandalus in the Skagerrak and the Norwegian Deep. The four species model was based on the primary predators identified by Jorgensen at al 2014, i.e., the three deep-sea species roundnose grenadier, velvet belly and blue whiting, plus the Atlantic cod. The nine species model added starry ray, hake and the gadoid species; haddock, whiting and saithe. Data from the Norwegian bottom trawl survey for northern shrimp (NO-SH) and from the International bottom trawl survey in the North Sea (NS-IBTS Q1/Q3) was used in the analyses.

Standardized catch per unite effort (*cpue*) was modelled for single species using a delta-gam approach. The combined biomass indices was calculated by adding up the single species *cpue* for the four and nine species, respectively. Predation indices was calculated from the single species *cpue* by applying species specific weights based on pandalus consumption before aggregation. We also calculated the combined indices directly from the index of predator biomass in the Norwegian shrimp survey for comparison.

Comparison of biomass and predation indices show that the choice of predator composition has an appreciable impact on the index, but that the weighting of predator to get to the predation index contributes little to the final variation. A caveat here is that we used fixed weights for the entire time series. Prey preferences may vary depending of the size of the pandalus stock, the availability of other, potentially preferred prey and the predator's size distribution. Unfortunately time series of prey preferences of predators in the Skagerrak and Norwegian deep are not available.

In general, the four species index and the nine species index showed similar trends, but with some notable differences. In 2001 and 2002 the index of the nine species model reached its highest values in the time series while at the same time the four species index was decreasing indicating a lower than average predator biomass. From 2011 to 2013 both indices showed a step increase, but while the nine species index started to decline immediately after, the four species index peaked again in 2015. The difference between the models probably relates to the species composition. The three deep sea species used in the models is mainly caught at depths below 250 meters which is the depth limit for the NS-IBTS. In practice this means that the four species index is a modelled shrimp survey index, with added data on Atlantic cod from the NS-IBTS. In 2001 and 2002 the cpue index for haddock and whiting peaked, and the cpue of saithe was also going up. None of these species are included in the 4 species index. In 2015 and 2016, the cod biomass reaches its highest values in the time series which is reflected in the 4 species index where cod is the only gadoid and the roundnose grenadier stock is very low. But in the 9-species index the peak in cod biomass is hidden by other gadoids, contributing to the combined index.

The modelling exercise indicates that it is worth combining data from the various surveys to generate a single biomass or predation index for the Norwegian Deep / Skagerrak pandalus population (NDSK). But further work is required before the approach can be implemented in assessment. First, an evaluation of uncertainties in the model is needed. Second, model simplification and data pruning should be evaluated to improve model fit and prediction. Finally, the composition of predators going into the index needs to be settled.

References

Albert, O.T. 1994b. Ecology of haddock (Melanogrammus aeglefinus L.) in the Norwegian Deep. ICES journal of Marine Science 51: 31-44.

Bergstad, O.A. 1990. Distribution, population structure, growth and reproduction of the roundnose grenadier Coryphaenoides rupestris (Pisces:Macrouridae) in the deep waters of the Skagerrak. Marine Biology 107: 25–39.

Bergstad, O. A. 1991. Distribution and trophic ecology of some gadoid fish of the Norwegian Deep. 1. Accounts of individual species. Sarsia 75: 269-313.

Bergstad, O.A., Wik, Å.D., Hildre, Ø. 2003. Predator-Prey relationships and Food Sources of the Skagerrak Deep-Water Fish assemblage. J. Northw. Atl. Fish. Sci. 31: 165-180.

EMODnet Bathymetry Consortium (2018): EMODnet Digital Bathymetry (DTM). https://doi.org/10.12770/18ff0d48-b203-4a65-94a9-5fd8b0ec35f6

ICES 2005. Report of the Pandalus assessment working group, 27 October – 5 November 2004. ICES C.M. 2005/ACFM:05, 74 p.

ICES 2006. Pandalus assessment working group report (WGPAND), 26 October – 3 November 2005. ICES CM 2006/ACFM:10. 60 pp.

ICES Stomach Dataset 2010, ICES, Copenhagen

ICES. 2020. Manual for the North Sea International Bottom Trawl Surveys. Series of ICES Survey Protocols SISP 10-IBTS 10, Revision 11. 102 pp. http://doi.org/10.17895/ices.pub.7562

ICES. 2021a. Workshop on the production of swept-area estimates for all hauls in DATRAS for biodiversity assessments (WKSAE-DATRAS). ICES Scientific Reports. 3:74. 77 pp. https://doi.org/10.17895/ices.pub.8232

ICES. 2021b. Workshop on the production of abundance estimates for sensitive species (WKABSENS). ICES Scientific Reports. 3:96. 128 pp. https://doi.org/10.17895/ices.pub.8299

Jørgensen, M. et al. 2014. Introducing time-varying natural mortality in the length-based assessment model for the Pandalus borealis stock in ICES Div. IIIa and IVa east.

Maunder, M. N., and Punt, A. E. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research, 70: 141–159.

Meadhbh Moriarty, Suresh A Sethi, Debbi Pedreschi, T Scott Smeltz, Chris McGonigle, Bradley P Harris, Nathan Wolf, Simon P R Greenstreet, Combining fisheries surveys to inform marine species distribution modelling. ICES journal of Marine Science 77: 539–552

Skjæraasen, J. E. and Bergstad, O.A. 2000. Distribution and feeding ecology of Raja radiata in the northeastern North Sea and Skagerrak (Norwegian Deep). ICES journal of Marine Science 57:1249-1260.

Skorda, E. T. 2018. Stomach sampling and analyses of shrimp predators in Skagerrak. Master's Thesis. Aquatic Science and Technology, DTU Aqua.

Søvik, G. and Thangstad, TH. 2021. Results of the Norwegian Bottom Trawl Survey for Northern Shrimp (*Pandalus borealis*) in Skagerrak and the Norwegian Deep (ICES Divisions 3.a and 4.a East) in 2021. NAFO SCR Doc. 21/001

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73(1):3-36

WD11 Estimating weight changes of processed shrimp catches in fisheries for northern shrimp (*Pandalus borealis*) in Skagerrak, Kattegat and the Norwegian Deep (ICES divisions 3.a and 4.a East)

Guldborg Søvik[1], Mikaela Bergenius Nord[2], Ole Ritzau Eigaard[3], Katja Norén[2], Mats Ulmestrand[2]

[1]Norwegian Institute of Marine Research

[2]Swedish University of Agricultural Sciences, SLU

[3]DTU Aqua, Technical University of Denmark

Introduction

In the northern shrimp (*Pandalus borealis*) fishery in the Skagerrak and Norwegian Deep (ICES division 27.3.a and the eastern part of division 27.4.a), Norwegian and Swedish fishers traditionally have sorted the catch onboard in size fractions, and the catch fraction consisting of the larger females have been boiled onboard to be landed fresh, fetching high prizes (ICES 2021a). The Danish shrimp fishers started boiling some of the larger shrimp onboard in the beginning of the 2000s and land the boiled catch primarily in Sweden.

As shrimp lose weight when boiled, the ICES shrimp working group (NIPAG) has for many years corrected the commercial landings numbers from the three countries, i.e. the boiled fraction of the landings has been multiplied by a factor of 1.13, to obtain fresh weight (round weight) (ICES 2021a). The correction factor has been provided by Swedish scientists, but NIPAG has not been able to obtain information on when or by whom it was estimated. An FAO Fisheries Circular (2000) gives correction factors for boiled catches of *P. borealis* and *Pandalus ssp* obtained from different countries: 1.15 (Sweden: boiled, fresh), 1.10 (Canada: boiled, frozen) and 1.20 (Russia: boiled, frozen). Correction factors have been provided also for frozen landings (Table 1), but as shrimp catches from the Skagerrak and Norwegian Deep primarily are landed fresh, only the factor 1.13 has been used by NIPAG. Norwegian landings back to 2000 have been corrected (Ulmestrand et al. 2015). There are no Norwegian data on boiled/raw landings fractions before 2000. Swedish boiled landings have been corrected for all years. Danish boiled landings back to 2001 have been corrected (when the Danish fishermen started boiling larger shrimp onboard).

Weight loss of shrimp following boiling has been studied by several authors (Benjakul et al. 2008, Jantakoson et al. 2012, Lascorz et al. 2016, Manheem et al. 2012, 2013, Martínez-Alvarez et al. 2009, Niamnuy et al. 2007, 2008). The studies show that boiling induces water loss, and that water loss is increased by increased temperature, boiling time, core temperature, and time at a specific core temperature. Water loss also increases with increased salinity in the boiling water. One article highlights that saline water has a higher boiling temperature, thus when salt water starts boiling the temperature is above 100 ^{o}C and a shorter boiling time is needed to reach a specific core temperature. Some of the articles also show that it is not only water that is lost from the shrimp during boiling but also proteins, and protein loss is increased with increasing boiling temperature and salinity. None of the studies tested the effect of cooling time or cooling method on weight loss as samples were generally rapidly cooled with ice, and measured shortly after.

Norwegian shrimp fishers have for some years pointed out that a large fraction of the Norwegian fleet cool the boiled shrimp with water, not in air, and that cooling with water does not lead to weight loss. According to *Fiskerlaget Sør* (the southern branch of the Norwegian Fishermen's Association), onboard the larger Norwegian vessels, the boiled shrimp are now either sprayed with cold water or soaked in cold water after boiling. Cooling in air was earlier the standard method, but presently this is done primarily on the smaller vessels.

To investigate the effect of different cooling methods on weight loss due to boiling, an experiment was carried out in 2017 during the annual Norwegian shrimp cruise in the Skagerrak and Norwegian Deep. The present document reports the results of this study. Furthermore, the industry was asked to document the cooling methods presently in use by the Norwegian shrimp trawlers in the region. These results are presented together with an overview of the Norwegian fleet and landings.

The document also presents results from recent Danish studies on weight gain of raw shrimp catches following storing on ice.

The results in this document are relevant for all three Scandinavian shrimp fisheries and landings data. The main goals of the study are 1) (re)calculate correction factors used for estimating round weight of landed shrimp, 2) appraise if there are reasons for recalculating the Norwegian time series of commercial shrimp landings going into the annual assessment, taking into account different cooling methods, and 3) conclude regarding future potential correction of boiled and/or raw landings. Finally, we briefly present plans for further studies on the topic of weight loss following boiling.

Table 1. Correction factors provided by Swedish scientists. The ICES shrimp working group, NIPAG, has only used the factor 1.13, to correct landed weight of boiled shrimp.

Conservation / handling of shrimp	Correction factor
Fresh, raw	1.0
Fresh, boiled	1.13
Frozen, raw	1.075
Frozen, boiled	1.164

Methods

Norwegian landings statistics

Data on official Norwegian shrimp landings were obtained from the Norwegian Directorate of Fisheries. Information on landings per vessel length category has been available since 2005. The number of vessels per length category is presented for all years since 1995. Vessels ≥ 11 m in length are considered “large” vessels.

Information on conservation of landings has been available from the Norwegian Directorate of Fisheries since 2000, given as “sea boiled”, “salt boiled”, or “fresh”. In 2011, the Directorate introduced a new category, “on ice”. The sales organization in Skagerrak, *Skagerakfisk SA*, has confirmed that since 2011, they have used “on ice” only for raw shrimp. The sales organization in southwestern Norway, *Rogaland Fiskesalgslag SA*, on the other hand, used “on ice” both for boiled and raw shrimp, resulting in ambiguous recordings in the official statistics. The proportion of boiled landings from the Norwegian Deep from the years 2011-2019 was therefore obtained directly from this sales organization. In 2015, shrimp sold through *Rogaland Fiskesalgslag SA* constituted 83 % of the total landed volume from division 4.a East. In 2020, the two sales organizations were merged, and the new sales organization, *Fiskehav*, uses “on ice” only for raw shrimp.

Information on correction of landings based on conservation method carried out by the Norwegian Directorate of Fisheries was obtained from the Directorate.

Cooling methods of Norwegian boiled shrimp catches

Information on how shrimp catches in the Norwegian fishery in the Skagerrak and Norwegian Deep are handled onboard was obtained from 1) information brochures on catch handling written and distributed to the fleet by *Informasjonsutvalget for reker* (the Norwegian Information Panel for Shrimp), last updated in 2015, 2) the sales organizations *Rogaland Fiskesalgslag*, *Skagerakfisk SA* and *Fiskehav*, 3) the secretariat of *Fiskerlaget Sør*, and 4) directly from a handful of shrimp fishers.

Experiment during shrimp cruise

An experiment for estimating weight loss after boiling of shrimp and testing out the effect of two different cooling methods was conducted in January 2017 on the annual shrimp cruise conducted by the Norwegian Institute of Marine Research (IMR) in the Skagerrak and Norwegian Deep, onboard the research vessel R/V *Kristine Bonnevie*.

Shrimp catches from some trawl stations were sorted into large and smaller shrimp in a sorting machine ("sollemaskin") (Figure 1). Samples of 1-2 kg were taken from the catch fraction containing large female shrimp, up to six samples from one single catch (Table 2). For a total of 27 shrimp samples the following was done:

The shrimp in the sample were counted, and the sample was weighed.

The samples were boiled individually in a shrimp boiler onboard for 4-5 minutes (Figure 1). Salt was added to the water which was boiling when the shrimp were added.

The samples were put individually in big plastic baskets and placed on deck (Figure 1) where they were cooled, either with water or in air.

Water: the samples were sprayed with cold sea water from a hose and left for 15 minutes for the water to drain off the shrimp.

Air: the samples were left to cool on deck for approximately 1 hour, after which the shrimp were lukewarm.

After 15 minutes (water cooling) or 1 hour (air cooling) the samples were weighed and then frozen individually at -20 °C.

The frozen samples were thawed, left for 15 minutes for the water to drain off, and then weighed again.

Factors were calculated as weight of treated sample divided by weight of raw sample.

Figure 1. Sorting machine ("sollemaskin") (left), photo: Guldborg Søvik. Boiling shrimp onboard R/V *Kristine Bonnevie* (middle); when the shrimp are floating on the surface they are finished, photo: Merete Kvalsund. Boiled shrimp in shrimp basket (right), photo: Heidi Gabrielsen.

Table 2. Overview of the trawl stations from which samples were taken: serial number, date, geographic position (latitude and longitude), air temperature during trawling (°C), number of samples from trawl haul, and cooling method (cold water or in air).

Serial number	Date	Position	Air temperature	Number of samples	Cooling method
22043	15.01.2017	58 09.94 N 006 31.73 E	-0.6	2	Water
22051	16.01.2017	57 40.28 N 006 28.49 E	4.2	4	Water
22053	16.01.2017	57 38.47 N 006 48.45 E	4.0	2	Water
22058	17.01.2017	57 30.94 N 007 27.34 E	5.2	3	Water
22062	17.01.2017	57 38.35 N 007 56.12 E	6.1	2	Air
22063	17.01.2017	57 36.08 N 007 57.02 E	6.5	5	Air
22068	18.01.2017	57 43.68 N 008 25.11 E	6.5	6	Air
22074	18.01.2017	57 57.15 N 009 17.86 E	6.6	3	Water

Danish experiments on raw shrimp on ice

In 2020, data were collected by DTU Aqua observers during three experiments onboard commercial vessels: *Sajoni* S486 (four days storage on ice, 24-29/10, 2020), *Skagerak* S107 (four days storage on ice, 31/10-5/11, 2020), and *North Sea* S255 (six days storage on ice, 8-14/11, 2020). In 2021, similar experimental data were produced during sea trials in September and November.

The shrimp for the experiments were kept cooled under normal commercial conditions in batches of 10 kg, which were weighed immediately upon storage and subsequently with 24 hours interval during the duration of the six trips in 2020 and 2021 (details in Danish in Appendix 3).

Results

Norwegian shrimp landings

The number of vessels in the Norwegian shrimp fleet in the Skagerrak and Norwegian Deep has decreased from around 400 in 1995 to presently around 200 (Figure 2). There has been a steady decrease in the number of vessels in the length categories < 10 m, 11-14.99 m, and 15-20.99 m, while the number in the length category 10-10.99 m increased until the mid-2000s and has since fluctuated without trend. The number of vessels ≥ 21 m has fluctuated without trend since 1995.

Landings from the larger vessels (≥ 21 m) have fluctuated without trend since 2005, between 1970 and 3500 tons (Appendix Table 1), but the proportion of landings from the two largest length groups have increased (Figure 3) and, in 2020, made up about 55 % of the total landings. The increase is explained mainly by the medium sized vessels (15-20.99 m) landing less.

The proportion of boiled landings has been larger in the Norwegian Deep compared with the Skagerrak through the whole time-series, except in 2020 (Figure 4, Appendix Table 2). This may reflect the population structure in the Norwegian Deep, with traditionally larger-sized shrimp as well as much fewer shrimp of the recruiting year class (1-year old shrimp) compared with in Skagerrak (Søvik and Thangstad 2021). The proportion of boiled landings (from uncorrected, official numbers) has since 2000 fluctuated between 43 and 77 % in the Norwegian Deep, and between 33 and 65 % in the Skagerrak (Appendix Table 2). The highest proportion in both regions was seen in 2010, the year with the lowest number of 1-year old shrimp observed so far in the Skagerrak and Norwegian Deep stock (ICES 2021a).

Correcting for boiling by NIPAG has implied that between 290 and 550 tons has been added to the official Norwegian landings numbers since 2000 (Appendix Table 3).

The Norwegian Directorate of Fisheries calculates round weight from product weight. In 2020, it was discovered that – due to poor communication – double correction of Norwegian boiled and frozen shrimp landings has been carried out since 2018 (by both IMR and the Directorate). The Directorate of Fisheries uses correction factors based on the results from the experiment on the IMR shrimp cruise in 2017 (Table 3). However, as Norwegian shrimp landings in the Skagerrak and Norwegian Deep primarily are landed fresh, correcting for freezing leads to minor weight increases. There is in fact no difference between product weight and round weight (in tons) in the official data in 2018-2020 (Appendix Table 3).

Table 3. Correction factors by year and product type (conservation method) used by the Norwegian Directorate of Fisheries for Norwegian shrimp landings. Information obtained from the Norwegian Directorate of Fisheries.

Years	Conservation method	Factor
2000-2017	All products	1.00
2018-present	Salt boiled, frozen	1.02
	Sea boiled, frozen	1.02
	Raw, frozen	1.02
	Salt boiled, fresh	1.00
	Sea boiled, fresh	1.00
	Other products	1.00

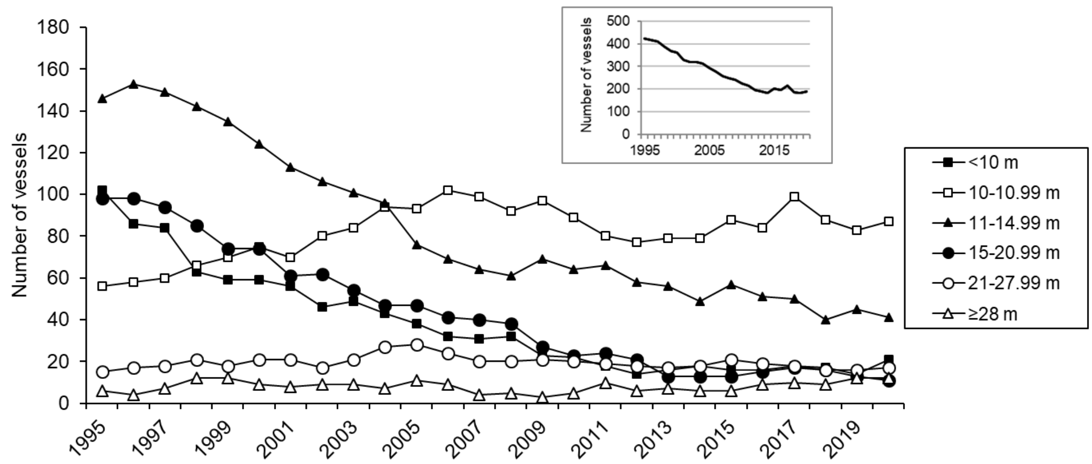


Figure 2. The number of vessels per length category in the Norwegian shrimp fleet in the Skagerrak and Norwegian Deep, 1995-2020 (all vessels landing shrimp are included). The inset figure shows the total number of vessels in the region per year. Data from the Norwegian Directorate of Fisheries.

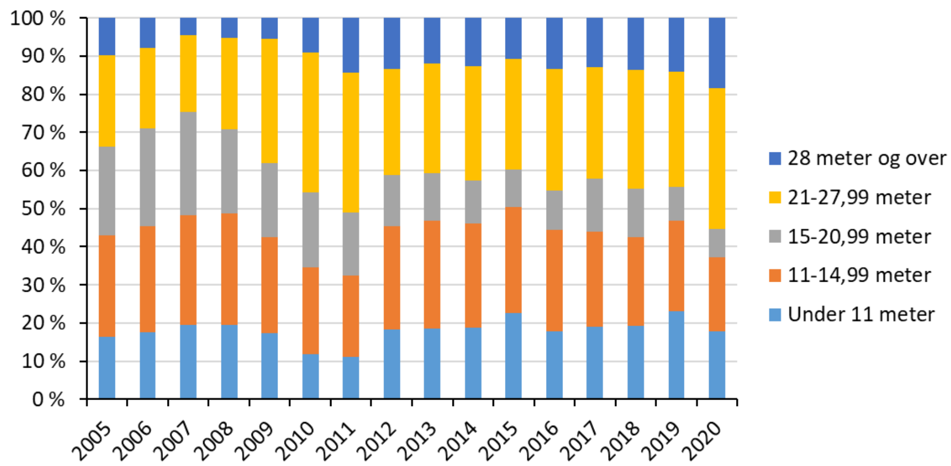


Figure 3. Norwegian shrimp landings from the Skagerrak and Norwegian Deep by vessel length category (% of total landings), 2005-2020. Data from the Norwegian Directorate of Fisheries.

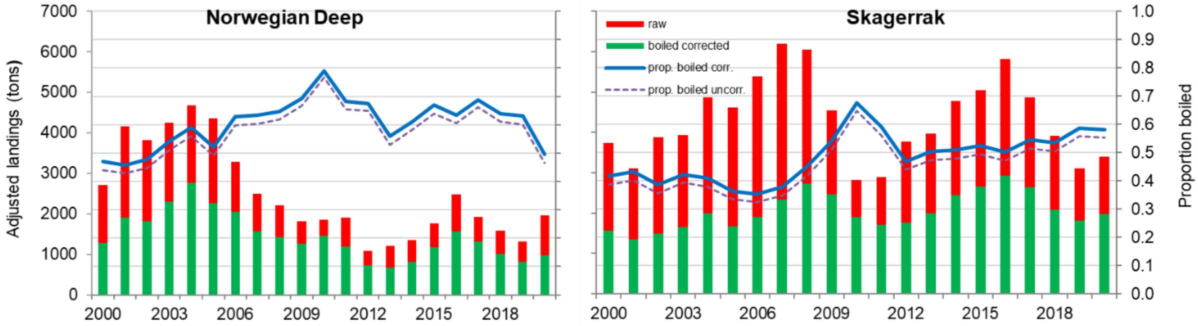


Figure 4. Norwegian boiled and raw shrimp landings from the Skagerrak (left) and Norwegian Deep (right), where the boiled fraction has been corrected by the factor 1.13 (Table 1), and annual proportions of boiled landings, from both corrected and uncorrected data. Data from the Norwegian Directorate of Fisheries and the sales organization *Rogaland Fiskesalgslag SA*.

Information from the Norwegian industry on catch handling

Informasjonsutvalget for reker (Information Panel for Shrimp)

According to information from *Informasjonsutvalget for reker* (Information Panel for Shrimp) from 2015 ([Lysbilde 1 (fiskehav.no)](fiskehav.no)), the salt content in the boiling water should be around 10 % (less than 1 kg per 10 liters of sea water). Shrimp are added to boiling water and should boil for 3-6 minutes depending on size. The boiled shrimp may be cooled either in water or in air. After cooling in water (sea water with ice), the shrimp (now holding 2 °C) are transferred to boxes where the water drains off. Thereafter, the shrimp are quickly packed in boxes and stored at temperatures between 0 and 2 °C. For cooling in air, the shrimp are spread out in boxes (half filled) and stored in a space with good ventilation. The shrimp should be cooled to a temperature of 2 °C within 2 hours.

Sales organizations

In 2017, both *Rogaland Fiskesalgslag SA* and *Skagerakfisk SA* confirmed that Norwegian shrimp vessels used both cooling methods for boiled shrimp, with water and in air. The small vessels (< 11 m) cooled in air, while the larger ones used water. In Skagerrak, a rough estimate indicated that approximately half the vessels cooled boiled landings in water, which amounted to roughly 70 % of the landings.

Fiskerlaget Sør (Norwegian Fishermen's Association)

According to *Fiskerlaget Sør*, fishermen operating in the Oslo fjord tend to deploy air cooling more than fishers along the rest of the coast, but the vessels spray the boiled fraction with cold water before air drying in order to quickly cool down the catch. Fishermen tell that the weight loss is less when the boiled shrimp are sprayed with cold water before cooling in air, compared with only air cooling. In this region, some large vessels also choose to air cool the shrimp in order to obtain better prizes on auctions in Sweden. In the cities Grimstad and Arendal, there are also several vessels that air cool the shrimp to obtain the right texture and shell colour. Fishermen in the west are more concerned about minimizing weight loss and ensuring high durability of the catch. Vessels that cool the boiled shrimp in brine use different equipment and techniques. Some cool the catch in freezing brine, especially larger vessels, while others cool the catch in sea water with extra salt added. With the exception of trawlers in the Oslo fjord region, most vessels are planning to shift to brine cooling in the near future.

Fiskerlaget Sør questioned a total of 243 Norwegian shrimp vessels about their landings in 2017-2019 and their preferred cooling method (Table 4). Several different methods exist for cooling boiled shrimp (see below for examples). In the survey, the different methods were categorized as either 1) spraying with cold sea water between 10 and 120 seconds, then air cooling / drying, or 2) cooling in cold brine. The results show that cooling method depends on the size of the vessel, with smaller vessels tending to cool the boiled shrimp in air, and the larger ones tending to cool in brine (Figure 5). As there are many more small than large vessels in the Norwegian shrimp fleet (Figure 2), more small vessels were interviewed. The large majority of the 243 vessels therefore reported that they cool their catch in air. However, as the large vessels land disproportionately more shrimp than the smaller ones, the total landings (in 2017-2019) from vessels cooling in brine was higher than the corresponding landings from vessels cooling in air (Figure 5); 46 % of the total landings from the 243 vessels was fished by trawlers cooling the boiled shrimp by spraying/air cooling, while 54 % was fished by trawlers cooling the shrimp in brine.

Table 4. Number of questioned shrimp trawlers by vessel length category in the survey conducted by *Fiskerlaget Sør* among their members, regarding method for cooling boiled shrimp.

< 11 m	11-14.99 m	15-20.99 m	21-27.99 m	≥ 28 m
121	62	22	25	13

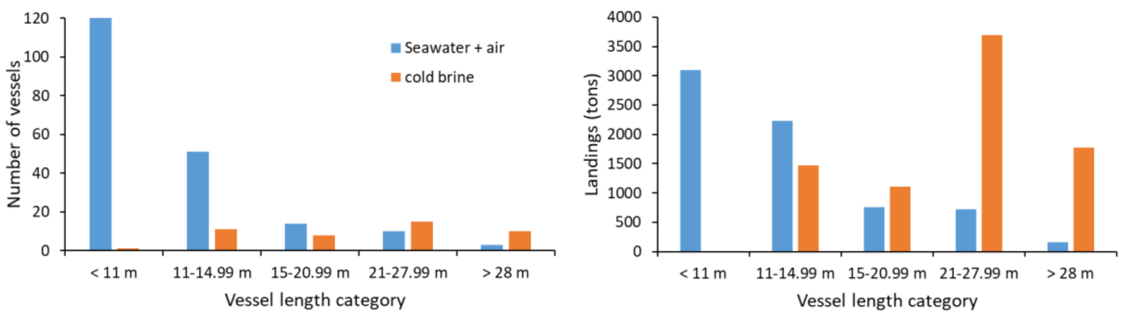


Figure 5. Information on cooling method from 243 interviewed Norwegian shrimp trawlers in the Skagerrak and Norwegian Deep: number of vessels per vessel length category and cooling method (left), and total landings (tons) in 2017-2019 per vessel length category and cooling method (right). The many different methods that exist for cooling boiled shrimp were categorized as either 1) sprayed with cold sea water in 10-120 seconds and then air cooled/dried, or 2) cooled in cold brine.

Norwegian shrimp fishers

Accounts from three different shrimp fishers serve to illustrate the different cooling methods in use by the Norwegian fleet. The skipper on a medium sized shrimp vessel (length 14.97 m) gave the following account in 2017 on how they boiled and cooled the shrimp onboard:

"The shrimp are boiled in sea water. The boiler contains about 600 liters of water. Approximately 100 kg salt is added to the water. The shrimp boil for about 4 minutes. They are thereafter placed in a container with cold sea water for rinsing and quick cooling before they are put in boxes for the water to drain off and further cooling for about 1 hour. Finally, they are packed in plastic and ice. If the vessel is fishing close to land and is on its way to shore, the shrimp are not packed in plastic and ice, but transferred to a cool storage room as soon as the vessel reaches the landing facility."

The ship owner of two larger trawlers (lengths 27.4 and 35.3 m) informed us in 2021 about their procedures onboard:

"The shrimp are boiled in 240 seconds before they are transferred directly to circulating sea water for 10-15 seconds. Thereafter, the shrimp are transferred to a mixture of brine and ice for two minutes."

A crew member on another medium sized shrimp trawler (length 12.78 m) informed us in 2021 about how they process the shrimp onboard:

“The shrimp are transferred directly to cooling tanks which are filled with sea water, where they remain between 10 and 15 minutes.”

This fisherman remarked that the shrimp probably lose some weight, but he could not estimate how many percent. He added that earlier it was almost “prohibited” to cool the shrimp with water, however, nowadays most fishers use this method, also the smallest vessels that don’t cool the shrimp in water tanks, but instead flush them in boxes after boiling.

Boiling experiment on R/V *Kristine Bonnevie*

Twenty-seven samples of shrimp were boiled and weighed, where 14 samples were cooled with water and 13 in air (Appendix Table 4).

Samples cooled with water gained on average 0.4 % weight, while the shrimp cooled in air lost on average 6.5 % weight. When frozen and thawed, the water-cooled and air-cooled samples lost on average 1.1 and 7.3 % weight respectively.

Table 5. Correction factors based on boiling experiment on R/V *Kristine Bonnevie* in 2017, mean with SD, where n = 14 for water cooled-samples, and n = 13 for air-cooled samples.

	Factor_boiled	Factor_boiled-frozen	Factor_boild-frozen-thawed
Water	1.00 (0.01)	1.00 (0.01)	1.01 (0.01)
Air	1.07 (0.02)	1.07 (0.04)	1.08 (0.02)

Danish experiment on raw shrimp on ice

In total six trials were conducted in 2020 and 2021 with scientific staff on commercial vessels to quantify the weight gain in raw shrimp during storage on ice, four trials lasted in four days, one in five days and one in six days. The study showed that the shrimps gained most weight the first day, increasing the weight by 7 % on average. In the following days, the shrimps added 1.4 % weight in average until the studies were terminated (Table 6, Figure 6). For Danish vessels, it is very rare that shrimp trips last more than four days.

Table 6. Weight changes (%) with time from the six different shrimp experiments.

Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
nov-20	8.3	8.6	9.3	12.8	13.2	14.6
nov-20	6.5	8.5	9.1	9.7		
okt-20	5.3	9.0	8.1	10.3		
sep-21	8.5	10.7	14.4	12.0		
sep-21	7.2	11.6	11.6	12.5		
nov-21	5.9	8.6	9.9	10.0	11.4	
average	7.0	9.5	10.4	11.2	12.3	14.6
SD	1.3	1.3	2.3	1.4		

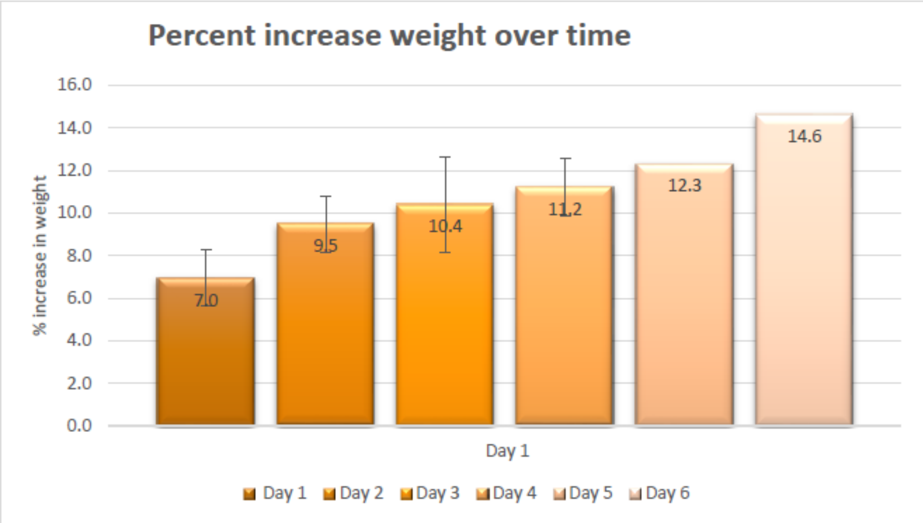


Figure 6. Average shrimp-weight changes with uncertainties (SD) (day 1-4) from the six experiments. Day 5 had data from two studies and day 6 from only one study.

Discussion

Correction factor for boiled shrimp landings

The Norwegian boiling experiment and the literature show that weight loss following boiling of shrimp catches depends on many factors: 1) the size of the shrimp that are boiled as this affects the time to reach a specific core temperature, 2) boiling time, 3) core temperature that is reached and the time this temperature is kept, 4) salinity in the boiling water, as higher salinity increases boiling temperature and the amount of proteins lost, and finally 5) cooling method and likely length of cooling. The Norwegian questionnaire survey shows that the handling of shrimp catches onboard varies between vessels, with a tendency of smaller vessels cooling the boiled catch fraction in air, while larger vessels cool the boiled shrimp in water or brine (but with exceptions). The Swedish fleet also consists of both large and small vessels while the Danish fleet consists only of large vessels (Ulmestrand et al. 2016). Differences in cooling method deployed may be expected between the three fleets, and may depend just as much on national market and consumer preferences as vessel length. One single correction factor for all boiled landings from the three countries might therefore not be correct.

Weight loss of shrimp following boiling has been studied by several authors (Benjakul et al. 2008, Jantakoson et al. 2012, Lascorz et al. 2016, Manheem et al. 2012, 2013, Martínez-Alvarez et al. 2009, Niamnuy et al. 2007, 2008). However, none of these studies can be used for deriving a correction factor useful for our purpose (converting the weight of fresh boiled landings to round weight), for several reasons:

Different species were studied

Individual shrimp have different sizes which affects boiling time for reaching a specific core temperature

Both boiling and steam boiling was used

Peeled, unpeeled and shrimp without head were used

Both fresh water and different salinities were tested

Fresh and sometimes frozen shrimp were used

Sometimes boiling time was reported, other times core temperature was tested

Different cooling methods were not tested

The estimated correction factors from the experiment in 2017 (Table 5) differed from the one(s) presently in use by NIPAG. Without knowledge on how the NIPAG correction factors were derived, it is impossible to try to explain the difference. The estimated factors from the experiment onboard R/V *Kristine Bonnevie* may be underestimates. If the boiled shrimp cooled in air had been left to cool for longer than 1 hour, the weight loss might have been larger. To arrive at more robust results, the experiment should be repeated under realistic circumstances onboard commercial shrimp vessels, preferably also conducted under different environmental conditions (summer and winter air temperatures) (see below).

Correction factor for raw shrimp landings

The results from the Danish experiments in 2020 and 2021 with raw shrimp demonstrate a significant weight-gain in raw shrimp catches, likely due to the shrimp taking up water during storage and cooling on ice onboard. This implies that the official landings data for raw shrimp represent an overestimate and should be down-scaled with a correction factor.

Historic Norwegian landings data

It is not possible to document cooling method on each and every Norwegian shrimp vessel, and it must be assumed that all vessels ≥ 11 m cool boiled shrimp in water, while all vessels < 11 m cool boiled shrimp in air. The fleet segment < 11 m is the largest in number, but these vessels only land around 20 % of the Norwegian shrimp landings from the Skagerrak and Norwegian Deep. The cooling methods onboard Norwegian shrimp trawlers have likely gradually developed over the years, from a few vessels cooling with water, to presently most larger vessels cooling with water, but it is impossible to document the use of different cooling methods by year. A pragmatic approach is therefore to not change the historic Norwegian landings that have been reported to NIPAG (except for using the final numbers instead of the preliminary ones (Appendix Table 3)).

Conclusions and further studies

The literature review and results from the Danish experiments on weight changes of raw shrimp landings show that the issue of weight changes in shrimp catches following handling onboard is more complicated than first anticipated. Preliminary conclusions show that landings numbers from the two landed catch fractions probably both should be corrected by suitable correction factors to obtain round weight, downscaling the weight of the raw shrimp landings and upscaling the weight of the boiled shrimp landings (possibly with inter-country differences in correction). The different correction procedures should be considered together, and as new experiments are planned and data from newly carried out experiments remain to be analyzed (see below), NIPAG will await these before final conclusions are made regarding the most appropriate methods for correcting landings of shrimp to reflect fresh weight.

Swedish study

The Swedish and Norwegian, and some of the Danish fishers, sort the catch onboard into three size fractions: shrimp to be boiled, shrimp to be landed raw and "lus" (small shrimp that are either landed separately or discarded). As mentioned above, shrimp lose weight when they are boiled, and the boiled landings have for many years been corrected by 1.13 to reflect fresh weight before these are included in the stock assessment. Based on the results in this working document, it seems likely that boiling and cooling methods vary between fishers in the three countries, and that the methods also vary between vessels within countries. A Swedish project was therefore carried out during autumn 2021 to:

Conduct a literature review on the subject of weight loss of shrimp due to boiling and variables influencing the weight loss.

Conduct a telephone survey among Swedish shrimp fishers to understand the different methods used when boiling, cooling, storing and weighing.

Conduct a pilot study onboard three Swedish shrimp boats during autumn 2021 to get an understanding of weight changes in shrimp due to the different methods used onboard.

From the telephone survey and pilot study design a larger study for 2022 to determine the weight change in shrimp. (The study is only preliminary funded yet.)

Norwegian-Russian study

The Norwegian Directorate of Fisheries has informed us that in summer 2022 a joint Norwegian-Russian scientific cruise is planned in the Barents Sea in order to estimate and harmonize factors for converting shrimp product weight to round weight. The parties have agreed to conduct measurements and estimation of correction factors for the following shrimp products: raw, frozen shrimp and boiled, frozen shrimp. The fresh (boiled and raw) samples will be weighed before freezing, making the results relevant for the landings data also in the Skagerrak and Norwegian Deep.

References

Benjakul, S., Visessanguan, W., Kijroongrojana, K. and Sriket, P. 2008. Effect of heating on physical properties and microstructure of black tiger shrimp (*Penaeus monodon*) and white shrimp (*Penaeus vannamei*) meats. International Journal of Food Science and Technology 43: 1066–1072.

FAO Fisheries Circular No. 847, Revision 1. 2000. Conversion factors. Landed weight to live weight. FIDI/C847 (Rev.1). ISSN 0429-9329. 192 pp.

ICES. 2021a. Joint NAFO\ICES Pandalus Assessment Working Group (NIPAG). ICES Scientific Reports. 3:22. 25 pp. https://doi.org/10.17895/ices.pub.7917

ICES 2021b. ICES Advice 2021 – pra.27.3a4a – https://doi.org/10.17895/ices.advice.7834

Jantakoson, T., Kijroongrojana, K. and Benjakul, S. 2012. Effect of high pressure and heat treatments on black tiger shrimp (*Penaeus monodon* Fabricius) muscle protein. International Aquatic Research 4:19. http://www.intaquares.com/content/4/1/19

Lascorz, D., Torella, E. Lyng, J.G. and Arroyo, C. 2016. The potential of ohmic heating as an alternative to steam for heat processing shrimps. Innovative Food Science and Emerging Technologies 37: 329–335.

Manheem, K., Benjakul, S., Kijroongrojana, K. and Visessanguan, W. 2012. The effect of heating conditions on polyphenol oxidase, proteases and melanosis in pre-cooked Pacific white shrimp during refrigerated storage. Food Chemistry 131: 1370–1375.

Manheem, K., Benjakul, S., Kijroongrojana, K. Faithong, N. and Visessanguan, W. 2013. Effect of pre-cooking times on enzymes, properties, and melanosis of Pacific white shrimp during refrigerated storage. International Aquatic Research 5:1. http://www.intaquares.com/content/5/1/1

Martínez-Alvarez, O., López-Caballero, M.E., Gómez-Guillén, M.C. and Montero, P. 2009. The effect of several cooking treatments on subsequent chilled storage of thawed deepwater pink shrimp (*Parapenaeus longirostris*) treated with different melanosis-inhibiting formulas. Food Science and Technology 42: 1335–1344.

Niamnuy, C., Devahastin, S. and Soponronnarit, S. 2007. Quality Changes of Shrimp during Boiling in Salt Solution. Journal of Food Science 72(5): 289-297.

Niamnuy, C., Devahastin, S. and Soponronnarit, S. 2008. Changes in protein compositions and their effects on physical changes of shrimp during boiling in salt solution. Food Chemistry 108: 165–175.

Søvik, G. and Thangstad, T. 2021. Results of the Norwegian Bottom Trawl Survey for Northern Shrimp (Pandalus borealis) in Skagerrak and the Norwegian Deep (ICES Divisions 3.a and 4.a east) in 2021. NAFO SCR Doc. 21/001, Serial No. N7157. 38 pp. https://www.nafo.int/Portals/0/PDFs/sc/2021/scr21-001.pdf

Ulmestrand, M., Bergenius, M., Eigaard, O., Søvik, G. and Munch-Petersen, S. 2016. The Northern shrimp (*Pandalus borealis*) Stock in Skagerrak and the Norwegian Deep (ICES Divisions IIIa and IVa East). NAFO SCR Doc. 16/056. 18 pp.

Appendices

Appendix 1. Data on Norwegian landings

Appendix Table 1. Landings (in tons) per vessel length category, and total, in 2005-2020. Data from the Norwegian Directorate of Fisheries.

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Under 11 meter	1369	1435	1585	1513	1028	509	494	835	904	1080	1435	1383	1195	984	954	895
11-14,99 meter	2198	2256	2347	2272	1489	977	950	1240	1372	1567	1758	2050	1565	1196	973	983
15-20,99 meter	1950	2095	2220	1711	1152	847	739	618	611	653	630	802	882	655	367	375
21-27,99 meter	1980	1717	1635	1864	1918	1582	1638	1262	1400	1717	1836	2473	1841	1593	1236	1859
28 meter og over	817	638	376	403	327	390	636	616	581	731	681	1037	812	704	581	925
Uoppgitt	260	75	74	23	88	31	25	7	5	265	1	2	10	6	2	4
Totalt	8574	8217	8237	7785	6002	4336	4482	4577	4873	6013	6341	7747	6304	5138	4113	5042

Appendix Table 2. Boiled proportion of the official Norwegian shrimp landings from respectively the Skagerrak and Norwegian Deep. Data from the Norwegian Directorate of Fisheries and the sales organization *Rogaland Fiskesalgslag SA*. Since 2018, the boiled and frozen landings (official numbers) have been corrected by the Directorate using the factors in Table 3.

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Skagerrak	0.39	0.40	0.36	0.39	0.38	0.33	0.33	0.35	0.42	0.51	0.65	0.56	0.44	0.47	0.48	0.49	0.47	0.51	0.50	0.56	0.55
Norw. Deep	0.44	0.43	0.45	0.51	0.56	0.49	0.60	0.60	0.62	0.67	0.77	0.60	0.18	0.14	0.10	0.10	0.17	0.23	0.21	0.22	0.46
Norw. Deep - data from sales organization												0.65	0.65	0.53	0.58	0.64	0.60	0.66	0.61	0.60	

Appendix Table 3. Overview of Norwegian landings data from the Skagerrak and Norwegian Deep, 2000-2020, per ICES division and total. Corrected total includes estimated weight loss due to boiling. Data reported to the ICES shrimp working group NIPAG (ICES 2021b) are preliminary (before final data were available); the other data in the table are final (except for 2020). The difference between NIPAG numbers and corrected total for 2001-2002 is due to the inclusion of landings from 4.a West (Fladen Ground) in the NIPAG numbers. Numbers for product weight and round weight are extracted from a different file format (LSS-files); the difference between these numbers and Total in 2015 is unknown.

Year	NIPAG	Corrected total	Total	3.a	4.a East	4.b	Weight loss due to boiling	Product weight	Round weight
2000	6442	6 442	6 116	3 554	2 562		326		
2001	7288	7 270	6 896	2 959	3 933	4	374		
2002	7713	7 703	7 321	3 709	3 612		382		
2003	8186	8 185	7 730	3 736	3 986	8	455		
2004	9548	9 548	9 002	4 638	4 360	4	546		
2005	8958	8 958	8 506	4 419	4 087		452	8 506	8 506
2006	8669	8 669	8 214	5 177	3 037		455	8 214	8 214
2007	8688	8 688	8 235	5 928	2 307		453	8 235	8 235
2008	8261	8 261	7 783	5 744	2 039		478	7 783	7 783
2009	6362	6 368	5 940	4 268	1 668	4	428	5 940	5 940
2010	4673	4 696	4 307	2 598	1 687	22	389	4 307	4 307
2011	4800	4 801	4 466	2 693	1 773		335	4 465	4 465
2012	4852	4 861	4 573	3 564	1 000	9	288	4 572	4 572
2013	5179	5 179	4 871	3 739	1 132		308	4 871	4 871
2014	6123	6 124	5 749	4 500	1 249		375	5 749	5 749
2015	6808	6 809	6 369	4 741	1 628		440	6 341	6 341
2016	8305	8 263	7 746	5 449	2 297		517	7 746	7 746

2017	6778	6 751	6 295	4 537	1 758		456	6 296	6 296
2018	5493	5 489	5 133	3 670	1 462	0	356	5 132	5 132
2019	4414	4 415	4 111	2 899	1 212		304	4 111	4 111
2020	5349	5 377	5 038	3 182	1 856		339	5 038	5 038

Appendix 2. Experiment onboard R/V *Kristine Bonnevie*

Appendix Table 4. Overview of data with ID and serial numbe for all shrimp samples from the boiling experiment on R/V *Kristine Bonnevie* in 2017: weight (W) (in gram) of the raw, boiled, frozen and thawed samples; number of shrimp per sample; factors (F) of the boiled, boiled-frozen, and boiled-frozen-thawed samples; number of shrimp per kilo in sample; and cooling method. Where the number of shrimp per sample is not an integer, the sample contained a headless body or a single head. The weight of the frozen shrimp for sample 21 may be a typo (shaded cell).

ID	Serial number	W_raw	W_boiled	W_frozen	W_thawed	No_sample	F_boil	F_boilfroz	F_boilfrozthaw	No_kg	Method
1	22043	1108	1130	1110	1093	138	0.981	0.998	1.014	124.5	Water
2	22043	1224	1221	1216	1215	141	1.002	1.007	1.007	115.2	Water
3	22051	1181.5	1206	1201	1194	165	0.980	0.984	0.990	139.7	Water
4	22051	1508	1493	1490	1485	203	1.010	1.012	1.015	134.6	Water
5	22051	1093	1102	1097	1085	146	0.992	0.996	1.007	133.6	Water
6	22051	1256	1272	1267	1268	151	0.987	0.991	0.991	120.2	Water
7	22053	1141	1135	1126	1109	134	1.005	1.013	1.029	117.4	Water
8	22053	1011	1003	1001	995	114	1.008	1.010	1.016	112.8	Water
9	22058	1196	1196	1186	1178	146	1.000	1.008	1.015	122.1	Water
10	22058	1285	1296	1290	1273	153	0.992	0.996	1.009	119.1	Water
11	22058	1428	1399	1390	1380	168.5	1.021	1.027	1.035	118.0	Water

12	22074	922.5	937	930.5	910.5	117.5	0.985	0.991	1.013	127.4	Water
13	22074	1149	1152.5	1154	1133	137.5	0.997	0.996	1.014	119.7	Water
14	22074	1173	1191	1190	1175	137	0.985	0.986	0.998	116.8	Water
15	22062	1755	1582	1578	1578	220	1.109	1.112	1.112	125.4	Air
16	22062	1685	1533	1534	1525	211	1.099	1.098	1.105	125.2	Air
17	22063	1072	1015	1009.5	1005	128	1.056	1.062	1.067	119.4	Air
18	22063	1019	951	945	945	125.5	1.072	1.078	1.078	123.2	Air
19	22063	1379	1285	1280	1275	178	1.073	1.077	1.082	129.1	Air
20	22063	1003	915	910.5	905.5	110	1.096	1.102	1.108	109.7	Air
21	22063	1143	1097	1194	1091	137	1.042	0.957	1.048	119.9	Air
22	22068	1047.5	980	976	972	133.5	1.069	1.073	1.078	127.4	Air
23	22068	1108.5	1080	1079.5	1059	144.5	1.026	1.027	1.047	130.4	Air
24	22068	1436	1341	1342	1325	175	1.071	1.070	1.084	121.9	Air
25	22068	1170.5	1107	1103	1098	147.5	1.057	1.061	1.066	126.0	Air
26	22068	1195	1130	1132	1115	152	1.058	1.056	1.072	127.2	Air
27	22068	1152.5	1070	1066	1061.5	142	1.077	1.081	1.086	123.2	Air

Appendix 3. Danish experiments with weight gain in raw shrimp in 2020 and 2021

MEMO

Til **Danish Fisheries Agancy**

Vedr. **Weight changes in Fish and Shellfish**

Fra **DTU Aqua**

29 november 2021J.nr. 20/1019489

Ref: MSP/JOSTOU

Request

DTU Aqua has been requested by the Danish Fisheries Agency to investigate weight changes in the commercial fishery (shrimps, saithe, cod, Nephrops, anglerfish, hake and plaice) from fishing time to llanding time.

Summary

DTU Aqua has conducted four studies (November 2020, Marts 2020, September 2021 and October 2021) to investigate if fish and shellfish are changing weight over time from the time they are caught tothe time they are landed. In our study, shrimps (Pandalus) gained in average 7% weight the first day and increased the weight by 1.4% in the following days. The reason for the weight gain in shellfish is believed to be ice and water absorbed from the fish boxes. Nephrops gained lesser in weight 0.5% after four days and 2.1% after a six days period. Anglerfish was the fish species showing the largest degree of weight loss. In a ten days period the angler fish loosed in average 7.3%, most in the first daywith 1.4%. For all other fish species investigated (saithe, cod, hake and plaice) the weight loss was below 2.5% in a 10 days period. For fish the hypothesis is that the weight loss is caused by water loss from the fish.

Background

The industry has highlighted that for some fish and shellfish species they weight at the catch time is different from the weight at the landing time (especially for longer trips). This has been a challenge as if the landed fish has been investigated by the control authorities the weights in land has been different from the weights at sea. To investigate if this claim could be correct a study was designed by DTU Aqua.

In the first study in October 2020 tree trips were conducted on three different shrimp vessels in the time period 25/10 to the 8/11. On two of the trips the shrimps (Pandalus) were measured every day ina 4 day period on the last trip the study was conducted in 6 days. On all vessels the shrimp samples were taken from the last haul on the trip by a scientific observer and sealed. However, as it was not possible to receive enough fish from all the wanted species to conduct the study the same way as withthe shrimps, fish were bought from the fisherman in the same harbour conducting a 1 day trip and landed in the night before they were measured first time. All the fish and shellfish were stored in a coldstore and was hold in a similar way as on a commercial vessel were the fish/ shellfish was stored in staked fish boxes with ice. Every day a scientific observer were weighting the samples without the ice in the cold store.

The first study was conducted in October / November 2020 (only shrimps), second in March (cod and saithe), third in September 2021 (shrimps, saithe, cod, Nephrops, anglerfish, hake and plaice) and the forth in November 2021 (shrimps, saithe, cod, Nephrops, anglerfish, hake and plaice).

Results

In total 6 trials were conducted with shrimps, four trials lasted in 4 days, one in 5 days and one in 6 days. The study showed that the shrimps gained most weight the first day increasing the weight by 7%in average. In the following days the shrimps added 1.4% in average until the studies were terminated (table 1, figure 1). For Danish vessels it is very rare that shrimp trips are lasting more than 4 days.

Table 1. Weight changes (%) in time from the 6 different shrimp studies.

days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
nov-20	8.3	8.6	9.3	12.8	13.2	14.6
nov-20	6.5	8.5	9.1	9.7		
nov-20	5.3	9.0	8.1	10.3		
sep-21	8.5	10.7	14.4	12.0		
sep-21	7.2	11.6	11.6	12.5		
nov-21	5.9	8.6	9.9	10.0	11.4	
average	7.0	9.5	10.4	11.2	12.3	14.6
SD	1.3	1.3	2.3	1.4		

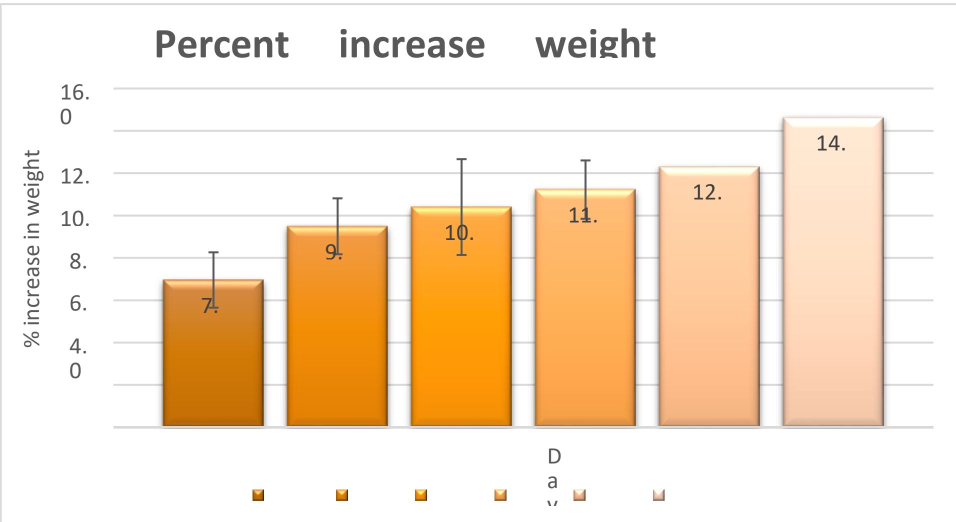
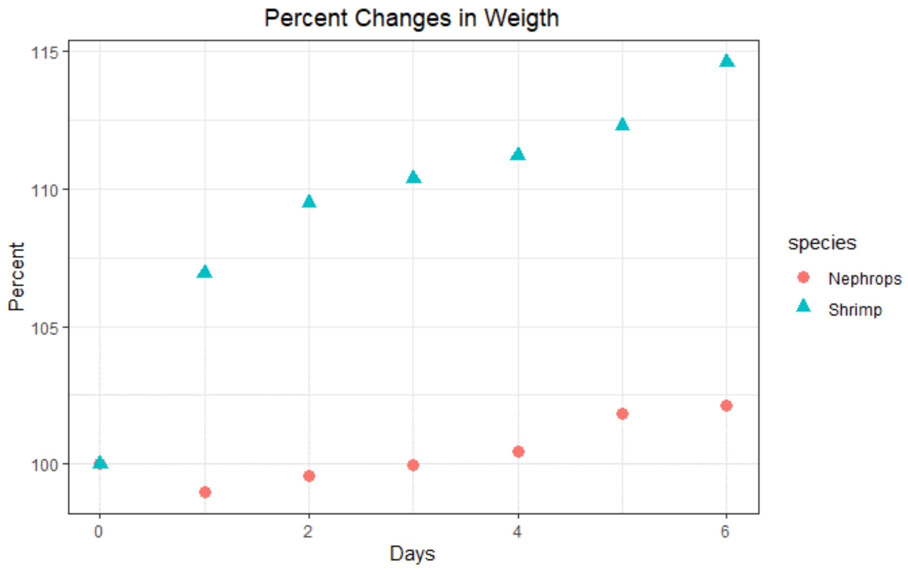


Figure 1. Average weight changes with uncertainties (SD) (day 1-4). Day 5 had twostudies and day 6 only one study.

Nephrops also being a shellfish could be expected to have a similar weight behavior as for shrimps. However, the weight gain was much smaller for the Nephrops compared with the shrimps with a small decrease of weight the first day. This could be caused by the way Nephrops are treated if they need to be stored for several days. The Nephrops are dipped in$NaHSO_3$ to slow down the decomposition. In total Nephrops gained 2.1% in a 6 days periodbut after four days the weight gain was only 0.5%. The experiment indicated that Neprops kept after day 6, stared to lose weight again, and at day 10 had lost 0.4 %.



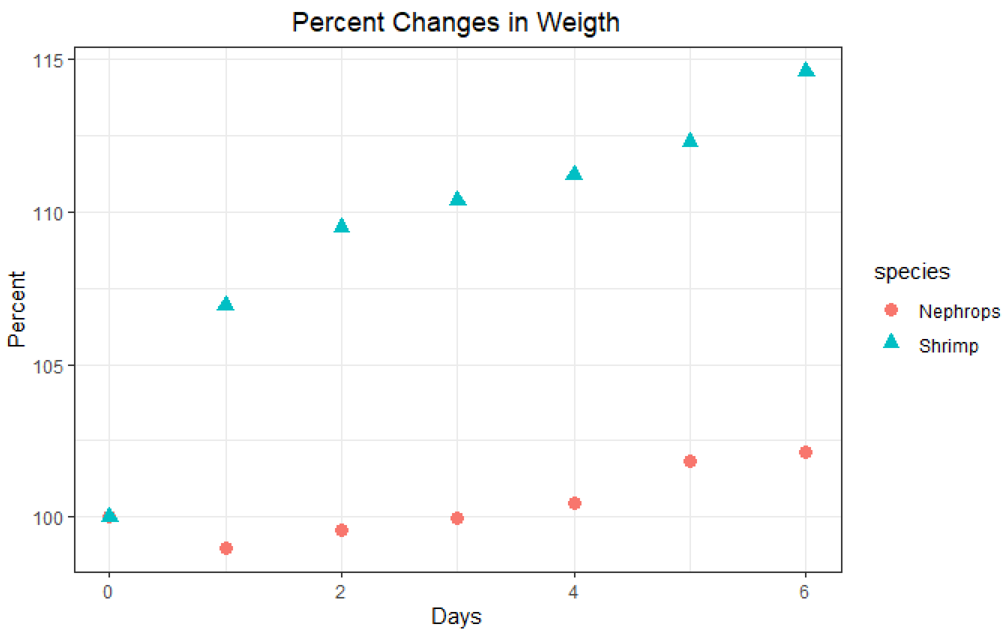


Figure 2. Average weight changes in a 6 days period for Nephrops and shrimps.

Table 2. Average weight changes (%) in time from the 6 different shrimp studies and 2 studieswith Nephrops.

Species	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
Shrimps	7.0	9.5	10.4	11.2	12.3	14.6
Nephrops	-1.0	-0.4	0.0	0.5	1.8	2.1

From the fish species investigated (saithe, cod, anglerfish, hake and plaice) the anglerfish were by farthe species showing the largest decrease in weight (7.3% in a 10 day period). The first day the weightloss in average was 1.4% and the following days 0.7%. For the remaining species, the changes in weight in a 10 days period were below 2.5% (saithe) and lesser for the other species (cod, hake and plaice) Figure 3 and table 3.

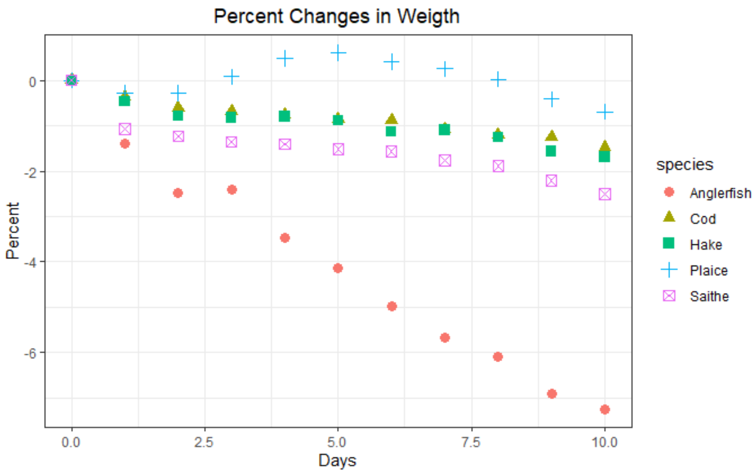


Figure 3. Average weight changes in a 10 day period for five different species. All trails wereconducted twice in September 2021 and November 2021 and three times for cod and saithe (March 2021 including).

Table 3. Average weight changes (%) in time from the five different fish species conducted in2 timer period for anglerfish, plaice and hake and three time period for cod and saithe).

DAYS	ANGLERFISH	COD	PLAICE	HAKE	SAITHE
1	-1,4	-0,4	-0,3	-0,5	-1,1
2	-2,5	-0,6	-0,3	-0,8	-1,2
3	-2,4	-0,7	0,1	-0,8	-1,4
4	-3,5	-0,8	0,5	-0,8	-1,4
5	-4,1	-0,9	0,6	-0,9	-1,5
6	-5,0	-0,9	0,4	-1,1	-1,6
7	-5,7	-1,1	0,3	-1,1	-1,8
8	-6,1	-1,2	0,0	-1,3	-1,9
9	-6,9	-1,2	-0,4	-1,6	-2,2
10	-7,3	-1,5	-0,7	-1,7	-2,5

WD12 Landings, discards, data coverage and quality from Denmark

By Ole Ritzau Eigaard, DTU aqua, Denmark

Landings, discards, data coverage and quality from Denmark

Description of the sources and methodology used to provide Danish input data on the *Pandalus borealis* stock in Skagerrak/Kattegat and Norwegian Deep (Pra.27.3a4a) for the 2022 ICES benchmark (WKPRAWN 2022)

Ole Ritzau Eigaard, DTU aqua, Denmark
Working document, ICES-WKPRAWN 2022

Shrimp landings-data from Denmark

Danish landings data for 1940-2020 was included in the benchmark assessment. This time series includes two periods with different sources of the landings data. From 1988 to 2020 the data of *Pandalus borealis* landings from 27.3a4a is based on EU logbooks and Danish sales slips, which in combination deliver high-confidence full coverage landings data at the fishing-trip level and at the spatial resolution of ICES rectangles. These logbook and sales slips data are hosted by the Danish Fisheries Agency (https://fiskeristyrelsen.dk/) and are freely accessible to DTU Aqua, where experts from the fisheries monitoring section have aggregated the data by year, quarter, and area (3a and 4a-east) to provide the landings data for 1988 to 2020 (Table 1 and Table 2). Prior to 1988, the benchmark input data (back to 1940) have been based on the official landings statistics in the data bases from ICES (https://www.ices.dk/data/Pages/default.aspx). The historical Danish landings data from this source are only available at the required spatial resolution (4a-east [Norwegian Deep] vs. 4a-west [Fladen Ground]) from 1970 onwards, even though both 4a-east and 4a-west have supported a targeted shrimp fishery since 1959. To enable a spatial segregation, the official combined landings from 4a for 1959 to 1969 have been split into landings from 4a-east and 4a-west, by applying a yearly area-proportion based on the data for subsequent three years (3-year average of 1970-1972). All the data prior to 1988 were split by quarter using average proportions as estimated from Danish national sources from 1988-1992 (5-years average). Landings prior to 2002 were not corrected for boiling, which coincides approx. with the years where the practice of on-board boiling was introduced in the Danish fishery.

Shrimp discard-data from Denmark

Danish discard data for 2009 to 2020 was included in the benchmark. This time series includes two periods with different resolution of the discard data. For both periods the discard information is derived from onboard sampling, which has been conducted by observers from DTU Aqua since 2009 with an approximate coverage of 1-2 trips and 5-7 hauls for each quarter (Table 3). This corresponds to about 1-2 % of all trips targeting *Pandalus borealis*. For the first four years of the program, 2009 to 2012, the samples of the landed fraction were not consistently split into raw and boiled shrimps and did not cover the quarters of each year very well, which was different from the sampling in the subsequent years (2013-2020). Instead, the 2009-2012 discard data for Denmark were taken from the ICES summary sheet and split by quarter and area using the average proportions as estimated with the higher-resolution data from 2013-2015 (3-years average). It is expected that for some of these four years (i.e., 2011 and 2012) high-resolution data can be established with further processing.

From 2013 to 2020 a total of 65 trips and 231 hauls were sampled with the objective to provide quarterly discard estimates for the fishery (Table 3). For the estimation of discards by weight and number in the fishery, the observer takes samples according to the below described catch process and haul-based sampling procedure.

For each haul the catch is lifted into the ship's pounder, from where it is processed along a sorting belt, which utilizes the differences in buoyancy between fish and shellfish together with serial sorting grids of different bar-spacings to split the catch into three fractions of shrimp: i) **small** shrimp (discarded shrimp with an average of 14 mm carapace-length), ii) **medium-sized** shrimp for landing raw (an average of 17 mm carapace-length), and iii) **large** shrimp for boiling on-board (an average of 21 mm carapace-length). An additional fraction consists of by-catch that can be landed, and typical species are saithe, cod, anglerfish, hake, and ling. The sorting process usually takes between 30 and 60 minutes, and subsequently the larger shrimp go through a longer boiling and air-cooling process (down to 3 degrees Celsius).

1. The catch of **small** shrimp is discarded continuously, in fairly stable amounts, together with other unwanted catch during the 30-60 minutes sorting process. This mixed discard slides down a ramp/quadratic metal channel to a conveyer belt and is returned to the sea though an open hatch in the ship side. The ramp/channel has a hatch in the bottom from which discard samples are taken at regular intervals during the sorting and discarding process. Normally 6-10 samples of approx. 10 kg each are taken and for each sample the opening time of the hatch is registered (typically between 30 and 60 seconds) to provide an estimate of shrimp and other discard in kg/hour. The full duration of the sorting/discarding process is also registered and the total amount (weight) of shrimp discard for the haul can be estimated after species-sorting and measuring the samples.
2. The catch of **medium-sized**, raw shrimp is weighed of in 20 kg boxes with ice and stored, and the total weight (number of boxes) is registered by haul by the observer.
3. The full catch of **large**, boiled shrimp is weighed of in 15 kg boxes after the cooling process and stored, and the total weight (number of boxes) is registered by haul by the observer.

The observer program provides approx. 5-7 samples per quarter and the individual weights of these samples are summed by catch-fraction before calculating a quarterly discard to landings ratio (the pooled weight of the discarded shrimp fraction to the pooled weight of the two landed fractions). Prior to the calculation of the quarterly discard percentages, the total weight of fraction 3 (large, boiled shrimp) is multiplied with a factor of 1.13 to correct for weight loss during the boiling and cooling process (ICES, 2021). The quarterly discard ratios were then applied to the total official landings (after correction for weight-loss due to on-board boiling and cooling) to provide total quarterly Danish discards of *Pandalus borealis* in 3a and 4a-east in weight for the years from 2013-2020. For the years 2009-2012, the quarterly discard estimates were based on data from the ICES summary sheet, split by quarter and area using the average proportions as estimated with the higher-resolution data from 2013-2015 (3-years average).

The fishing trips included in the Danish observer program are selected randomly, and consequently the spatial coverage is in proportion to the spatial distribution of the commercial fishery. As the Danish Pandalus fishery since 2013 almost exclusively has taken place in 3a (close to 95% of the landings), the observer data and the estimates of discard percentages are also almost exclusively from 3a (only 5 of 65 trips are from 4a-east). Consequently, it is not possible to provide area-specific discard estimates for 3a and 4a-east for more than a few quarters in a few years, and instead the observer data have been pooled for the two areas and the same discard percentages are applied to the total catches from both areas. The use of this approach is supported by the spatial distribution of the (limited) 4a-east fishery where the bulk of the landings are from ICES-rectangles close to the border to 3a.

Shrimp length-data from Denmark

The Danish length-data for the benchmark covers the period from 2013 to 2020. The observer program for discard sampling of the Danish fishery (described above), also forms the basis for the Danish sampling of length data from the *Pandalus borealis* stock in 27.3a4a. A subset of the hauls that are sampled to provide discard estimates are also sampled with the objective to provide estimates of the Length Frequency Distribution (LFD) of the commercial catches. In the period from 2013 to 2020 a total of 62 trips and 172 hauls were sampled for length measurements of Pandalus borealis in the laboratories of DTU Aqua (Table 3).

The length-data data are based on three separate samples from each of the fractions described above under the discard sampling; 1) small, discarded shrimp, 2) medium-sized shrimp landed raw, and 3) large shrimp boiled on-board before being landed. All three catch-fractions are sampled after sorting and raw before processing of any kind. The sample-size of each fraction in the catch of a haul is typically around 160-180 shrimp, with a little less for the small shrimp and a little more for the medium-sized shrimp. The shrimp are length-measured individually (carapace length with a precision of 1 mm) and subsequently weighed together by each 1 mm group. The weights of the 1 mm length-groups are then summed to provide total sample weight by fraction.

The observer program provides on average about 4-5 samples of shrimp lengths per quarter (Table 3) and the individual quarterly LFDs and weights of these samples are summed by each of the three catch-fractions before calculating a quarterly LFD per catch weight by fraction (the pooled LFD in absolute numbers to the pooled weight). These quarterly LFD/weight factors are raised to the total official catches (landings + discards) within each of the three fractions (discarded, medium-sized, and large shrimp). These fractions are then summed to provide total quarterly Danish catches of *Pandalus borealis* in 3a and 4a-east in absolute numbers per 1 mm length group (carapace length) for the years from 2013-2020.

Because of the way the commercial sorting process is designed and the customized sampling-protocol for obtaining discards in weight-percentages (as detailed under the discard sampling above), the calculations of the quarterly, pooled LFD raising-factors by weight (LFD/kg) involve a standardization of catch-weight to time. This means that there is an intermediate step, where the sample weight of each of the three catch-fractions is converted into kg catch per hour, based on information of the duration of both the haul and the sorting (discarding) process.

As described above the Danish observer-based sampling data are also almost exclusively from 3a (only 5 of 65 trips are from 4a-east) and consequently, it is not possible to provide area-specific LFDs for 3a and 4a-east for more than a few quarters in a few years. Instead, the LFD data have been pooled for the two areas and the same quarterly LFDs are raised to the total catches from both areas.

References

ICES. 2021. Joint NAFO\ICES Pandalus Assessment Working Group (NIPAG). ICES Scientific Reports. 3:22. 25 pp. https://doi.org/10.17895/ices.pub.7917

Table 1. Danish landings and discard data (tons) by year and quarter for Skagerrak/Kattegat (3a) for the period 1988-2020. The landings are corrected for boiling with a factor of 1.13 starting in quarter 4 in 2002. Total quarterly discards are calculated as a proportion of total landings (after boiling correction) based on quarterly data from observer trips.

	Quarter 1		Quarter 2		Quarter 3		Quarter 4	
Year	Land.	Disc.	Land.	Disc.	Land.	Disc.	Land.	Disc.
1988	546		862		496		355	
1989	318		909		923		379	
1990	340		566		851		476	
1991	685		847		946		763	
1992	982		765		869		683	
1993	369		585		1036		477	
1994	379		758		419		427	
1995	483		795		565		653	
1996	930		1110		694		923	
1997	782		987		895		937	
1998	673		870		852		539	
1999	127		397		537		344	
2000	408		225		640		610	
2001	180		338		407		263	
2002	324		392		599		631	
2003	611		602		658		734	
2004	500		891		851		821	
2005	484		716		705		609	
2006	778		721		686		688	
2007	559		427		748		636	
2008	309		425		794		645	
2009	496		494		427		563	
2010	372		280		301		227	
2011	216		359		438		307	
2012	195		285		421		380	
2013	359	74	596	41	629	34	405	29
2014	269	126	556	116	916	57	642	215
2015	528	64	634	44	801	50	564	31
2016	551	27	370	5	558	0	404	0
2017	453	90	463	92	555	9	584	0
2018	396	0	477	0	448	9	455	3
2019	335	15	430	33	660	27	390	1
2020	572	14	435	6	532	15	567	20

Table 2. Danish landings and discard data (tons) by year and quarter for Norwegian Deep (4a-east) for the period 1988-2020. The landings are corrected for boiling with a factor of 1.13 starting in quarter 4 in 2002. Total quarterly discards are calculated as a proportion of total landings (after boiling correction) based on quarterly data from observer trips.

	Quarter 1		Quarter 2		Quarter 3		Quarter 4	
Year	Land.	Disc.	Land.	Disc.	Land.	Disc.	Land.	Disc.
1988	743		413		44		14	
1989	237		246		259		105	
1990	73		219		91		151	
1991	99		164		36		46	
1992	285		293		46		44	
1993	285		183		20		21	
1994	94		74		11		1	
1995	27		61		42		18	
1996	187		125		37		50	
1997	156		206		17		64	
1998	297		86		90		5	
1999	497		96		138		7	
2000	207		230		112		63	
2001	687		174		51		0	
2002	206		213		159		34	
2003	453		220		40		15	
2004	725		136		21		2	
2005	386		81		19		2	
2006	178		43		6		12	
2007	82		13		3		0	
2008	39		49		14		1	
2009	116		81		29		18	
2010	51		33		11		25	
2011	140		84		45		4	
2012	115		33		19		8	
2013	35	7	2	0	1	0	0	0
2014	4	2	6	1	11	1	27	9
2015	58	7	6	0	62	4	37	2
2016	57	3	32	0	10	0	13	0
2017	73	15	0	0	13	0	17	0
2018	75	0	3	0	5	0	8	0
2019	40	2	0	0	134	5	59	0
2020	98	2	3	0	65	2	28	1

Table 3. Coverage of observer based on-board sampling, discard percentages in weight and number of length measurements from the Danish fishery from 2009 to 2020. Quarterly discard percentages are calculated as a proportion of total landings (after boiling correction) based on pooled data from the quarterly observer trips.

		Discard sampling (weights)			Length measurements		
Year	Quarter	Trips	Hauls	Discard (%)	Trips	Hauls	Measured
*2009**	*1*	-	-	-	-	-	-
	2	-	-	-	-	-	-
	3	*1*	*4*	*2.3*	*1*	*3*	*1096*
	4	*2*	*7*	*1.2*	*2*	*7*	*1776*
*2010**	*1*	*1*	*5*	*10.2*	*1*	*3*	*901*
	2	*1*	*3*	*2.7*	*1*	*2*	*942*
	3	*2*	*7*	*0.6*	*2*	*5*	*1861*
	4	-	-	-	-	-	-
*2011**	*1*	-	-	-	-	-	-
	2	*1*	*4*	*10.3*	*1*	*2*	*646*
	3	*3*	*6*	*6.1*	*3*	*6*	*2676*
	4	*1*	*4*	*5.1*	*1*	*3*	*1142*
*2012**	*1*	*2*	*6*	*7.9*	*2*	*6*	*3182*
	2	*2*	*4*	*3.4*	*2*	*4*	*1617*
	3	*1*	*4*	*3.7*	*1*	*4*	*1448*
	4	*1*	*3*	*9.9*	*1*	*3*	*1288*
2013	1	2	8	16.7	2	8	4026
	2	1	4	6.4	1	4	1840
	3	1	3	5.0	1	3	1204
	4	1	5	6.7	1	4	2468
2014	1	2	7	25.6	2	3	1899
	2	1	2	16.3	1	2	1064
	3	1	5	4.9	1	3	1862
	4	1	6	17.6	1	3	1938
2015	1	3	13	7.7	3	11	5602
	2	1	3	6.4	1	3	1643
	3	-	-	-	-	-	-
	4	1	4	5.0	1	4	2257
2016	1	1	3	4.6	1	3	1716
	2	2	9	1.9	2	6	2361
	3	2	9	0.0	2	5	2085
	4	2	7	0.0	1	3	1220
2017	1	-	-	-	-	-	-
	2	1	4	12.9	1	3	1800
	3	2	8	0.3	2	5	2038
	4	5	14	0.0	3	7	2912
2018	1	1	2	0.0	1	2	816
	2	2	9	0.0	2	6	2359
	3	2	7	1.8	2	6	2587
	4	2	8	0.4	2	6	2660
2019	1	2	8	2.7	2	8	3266
	2	3	9	5.6	3	7	3368
	3	2	10	1.9	2	6	2566
	4	2	2	0.2	2	2	1183
2020	1	2	8	1.9	2	6	3503
	2	2	4	1.3	2	4	1779
	3	3	7	2.4	3	7	3189
	4	3	11	2.7	3	9	2418

** For the years 2009 to 2012 the discard sampling did not consistently split the landed fraction into raw and boiled shrimps, and LFDs and discard percentages from these years are not included in the input data for the benchmark modelling. It is expected that appropriate data formats from some of these four years (i.e., 2011 and 2012) can be established with further processing.*