INTER-BENCHMARK PROCESS ON SANDEEL (*AMMODYTES* SPP.) IN AREA 2R (CENTRAL AND SOUTHERN NORTH SEA, DOGGER BANK), AND AREA 3R (SKAGERRAK, NORTHERN AND CENTRAL NORTH SEA) (IBPSANDEEL)

VOLUME 2 | ISSUE 11

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

The material in this report may be reused for non-commercial purposes using the recommended citation. ICES may only grant usage rights of information, data, images, graphs, etc. of which it has ownership. For other third-party material cited in this report, you must contact the original copyright holder for permission. For citation of datasets or use of data to be included in other databases, please refer to the latest ICES data policy on ICES website. All extracts must be acknowledged. For other reproduction requests please contact the General Secretary.

This document is the product of an expert group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the view of the Council.

ISSN number: 2618-1371 I © 2020 International Council for the Exploration of the Sea

ICES Scientific Reports

Volume 2 | Issue 11

INTER-BENCHMARK PROCESS ON SANDEEL (*AMMODYTES* SPP.) IN AREA 2R (CENTRAL AND SOUTHERN NORTH SEA, DOGGER BANK), AND AREA 3R (SKAGERRAK, NORTHERN AND CENTRAL NORTH SEA) (IBPSANDEEL)

Recommended format for purpose of citation:

ICES. 2020. Inter-benchmark process on Sandeel (*Ammodytes* spp.) in Area 2r (central and southern North Sea, Dogger Bank), and Area 3r (Skagerrak, northern and central North Sea) (IBPSandeel). ICES Scientific Reports. 2:11. 23 pp. http://doi.org/10.17895/ices.pub.5553

Editor

Valerio Bartolino

Authors

Valerio Bartolino • Mikael van Deurs • Espen Johnson • Claus Reedtz Sparrevohn • Anna Rindorf • Pia Schuchert • Nicola Walker

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

Contents

i Executive summary

The inter-benchmark IBPSandeel evaluated the use of density-dependent survey catchability to reduce retrospective patterns in the assessment of sandeel in area 2r and 3r. IBPSandeel concluded that the method proposed is appropriate and can be applied to provide advice on these two stocks.

The sandeel advice on fishing opportunities is highly influenced by good incoming year classes. High uncertainty is usually associated with estimation of high recruitments, which justified the implementation of an F_{cap} strategy. However, sandeel assessments in 2r and 3r have also strong retrospective patterns in recruitment (i.e. the model has a tendency to overestimate recruitment in the terminal year) which are not properly accounted by the current F_{cap}.

Sandeel recruitment in 2019 is estimated to be high throughout the entire North Sea, including areas 2r and 3r. For this reason, an adjustment to the assessment model was proposed to provide more reliable estimates of recruitment in the terminal year. A power function was implemented in the assessment of sandeel 2r and 3r to capture density-dependency catchability of age 0 fish in the survey. The adjusted models provided downward correction of the terminal year recruitment, which is considered more reliable as suggested by the reduced Mohn's rho statistic, while estimates of stock dynamics (SSB and R) remained highly consistent with the previous assessment.

ii Expert group information

Expert group name	Inter-benchmark process on Sandeel (*Ammodytes* spp.) in Area 2r (central and southern North Sea, Dogger Bank), and Area 3r (Skagerrak, northern and central North Sea) (IBPSandeel 2020)
Expert group cycle	Annual
Year cycle started	2020
Reporting year in cycle	1/1
Chair(s)	Valerio Bartolino, Sweden
Meeting venue and dates	29 January 2020, by correspondence, eight participants

1 Terms of Reference

IBPSandeel (*Ammodytes* spp) in Area 2r (Skagerrak, central and southern North Sea), and Area 3r (Skagerrak, northern and central North Sea)

Inter-benchmark process (IBP) on sandeel (Ammodytes spp.) in Area 2r (central and southern North Sea, Dogger Bank), and Area 3r (Skagerrak, northern and central North Sea), chaired by Valerio Bartolino, Sweden, and reviewed by external experts Pia Schuchert, UK and Nicola Walker, UK will be established and meet by correspondence on the 29 January 2020 to:

a) **Evaluate if the adapted assessment method of using density-dependent survey catchability to reduce retrospective bias in the assessment of sandeel in area 2r and 3r, will provide a more consistent forecast and if it is considered an improvement on the existing method.**

Stocks	Stock leader
Sandeel (*Ammodytes spp*.) in Divisions 4.b and 4.c, and Subdivision 20, Sandeel Area 2r (Skagerrak, central and southern North Sea)	Mikael van Deurs
Sandeel (*Ammodytes spp*.) in Divisions 4.a and 4.b, and Subdivision 20, Sandeel Area 3r (Skagerrak, northern and central North Sea)	Mikael van Deurs

The Benchmark Workshop will report by 31 January 2020 for the attention of ACOM.

2 Working document for use in inter-benchmark process (next benchmark has been scheduled for 2021, fall)

2.1 The use of density-dependent survey catchability as a method to reduce retrospective bias in the assessment of sandeel in area 2r and 3r

During the sandeel assessment and advice meeting (held on 22–24 January 2020 at the ICES HQ as part of HAWG), a retrospective bias in area 2r was observed by the expert group in both recruitment (R) and spawning–stock biomass (SSB) (i.e. Mohn´s Rho >0.5). In particular, a severe retrospective bias in recruitment can potentially be a problem in relation to the forecasted TAC (i.e. the TAC relies largely on the most recent age-0 survey index). In addition, SSB is below B_{lim} in 2020 and was close to B_{lim} in 2019. Since 2004, SSB has been below B_{pa} in 16 out of 17 years, and below B_{lim} in 14 out of 17 years in area 2r.

The Chair of the group and the stock assessor participated in ICES WKFORBIAS meeting in 2019, designated to come up with recommendation for these kinds of situations. However, since the recommendations are not yet ready, the group attended a preliminary decision tree (from WKFORBIAS) provided by ICES. According to the decision three, a severe retro pattern requires immediate action: (1) Identify the problem and make an adjustment of the model settings to solve the problem, or (2) adjust the forecasted TAC down, or (3) move the stock to category 3, as the analytical assessment cannot be used to provide advice as it is. The group therefore identified a model adjustment, which involved adding density-dependent catchability to the survey index-stock–number relationship for age-0.

In the SMS stock assessment model, survey index and stock numbers are modelled using the following equation: $Ln(I)=ln(qN^p)$, where I is survey CPUE (for a given age), N is the survey-estimated stock–numbers (at a given age), q is catchability, and p is the exponent of the power function. In the default model, p is forced to be 1, whereas, in the model with density-dependent catchability, p is estimated as a free parameter in the assessment model (i.e. the relationship between index and stock numbers is determined by a combination of q and p). If catchability increases with increasing densities, p should be >1 and vice versa. In Figures 1 and 9 (area 2r and 3r, respectively), it is seen that the slope is >1 if a linear regression is fitted to $ln(N_{age0})$ vs. $ln(Index_{age0})$, indicating increased catchability at high densities. Note also that a study from 2013 showed that the catchability of the sandeel dredge (same as used in the sandeel survey) increased with increasing densities (see Johnsen, E., and Harbitz, A. (2013)). Small-scale spatial structuring of burrowed sandeels and the catching properties of the dredge. ICES Journal of Marine Science, 70(2), 379–386). The power model approach is also used for North Sea sprat and was benchmarked in 2018 and presented at the WKFORBIAS meeting.

The addition of density-dependent catchability reduced the retrospective bias in recruitment and SSB in area 2r. The five-year Mohn´s Rho value were only slightly reduced. However, it can be seen from the individual peels (see Figure 5) that Mohn´s Rho is above 0.3 for recruitment, only because of one single retro-value, which belongs to the oldest of the peels in the time-series used. Therefore, the Mohn´s Rho to be calculated in 2021 is likely to be below 0.3. The pattern seen for SSB is similar, but in this case are the two oldest peels which inflate the Mohn's rho estimation.

As explained in the benchmark report, exploitation pattern is estimated in time blocks and the last of these blocks is based on data from 2010 and onwards, which may partially explain the very poor retro values among the oldest peels. In addition, when looking at the individual retro peels for recruitment (Figure 5), it is evident that the large retro corresponding to the observed the last high recruitment (2016) has disappeared after introducing density-dependent catchability, indicating that such adjustment is specifically addressing the model tendency to overestimate large recruitments.

If introducing density-dependent catchability to the model causes major changes to the model output (e.g. the stock–recruitment plot constituting the basis for B_{lim}) or reduces model quality, this would pose a problem. However, this was not the case. The position of the S–R pairs in the plot is reasonably stable and all the year's position at the two sides of B_{lim} (vertical dotted line in Figures 7 and 14) consistently between the two models. In Figures 2–8 (area 2r) and Figures 10–16 (area 3r), it is seen that the model as such reconstructs recruitment and SSB time-series consistent with the last year's assessment, and the different diagnostics, such as AIC, survey and catch CV, survey residuals, either remain the same or are improved. Only exception is the reduction of recruitment in the terminal year (and to some extent large historic recruitment values within the time period of the dredge survey). Reduced recruitment in the terminal year will in turn affects the catch advice, since the catch advice is largely determined by how many age-1 fish we predict will enter the fishery in the coming season (which is predicted based on the age 0 from the dredge survey). Taking area 2r as an example, recruitment in 2019 (the terminal year) was down-scaled by 26.7% when density-dependent catchability was implemented. The recruitment index for 2019 in area 2r was high (second highest in the survey time-series).To investigate what would happen when the recruitment index from the survey is low, we implemented density-dependent catchability in the last year assessment (the 2018 recruitment index was very low) and found only a 4% downscaling of recruitment. According to our expectations and intentions, this demonstrates that the effect of using a power model for the age 0 survey catchability affects recruitment estimates mostly when the recruitment index from the survey is high. Moreover, in the historic recruitment values (i.e. years before the terminal year), the change on average only 1% (incl. both positive and negative changes) and the largest change is a 7% reduction observed in 2016 (by far the largest recruitment index in the survey time-series). The explanation for this is that the historic recruitment values are informed by catch data and relies much less on the survey, whereas, recruitment in the terminal year relies solely on the survey (therefore, if recruitment is adjusted downwards when it is being informed by catch data, it tends to cause a retrospective pattern like the one we see here).

The working group also discussed if implementation of density-dependent catchability could be done without re-estimating F_{cap} and the biomass reference points. Since the current F_{cap} values are estimated in a simplified MSE (a full MSE model for sandeel does not currently exist, but is planned for the benchmark scheduled for 2021), stock number CVs in the terminal year are used in the calculations of F_{cap}. Hence, if adding density-dependence reduces the uncertainty (represented by the CVs) considerably, it could be argued that F_{cap} values need to be updated. However, Table 1 shows that the reduction in CV is much smaller than the interannual variability of the estimated CV in the terminal year. This suggests that F_{cap} should not be recalculated based on the change in the terminal year CV as it is not recalculated every year according to the ICES procedures. Given that B_{lim} is unchanged, and the terminal year CV falls within its interannual variability, the B_{pa} available from the last benchmark (ICES, 2016) is considered still valid for the adjusted model ($B_{pa} = B_{lim} \times \exp(\sigma \times 1.645)$, with σ estimated from the assessment uncertainty in the terminal year). Furthermore, updating the MSEs would require thorough expert group discussions about other MSE inputs, such as the parameters and assumptions used to predict future recruitment, weight-at-age and other stock metrics, which is not compatible with the present timeframe.

After having decided to propose to apply density-dependent catchability for area 2r, it was decided, that in order to be consistent, density-dependent catchability should be implemented to all areas where large retrospective biases in recruitment were observed (unless it reduces model quality or change the stock assessment model notably, e.g. the stock–recruitment plot constituting the basis for B_{lim}).

Area 3r also displayed a retrospective bias in recruitment (Mohn´s Rho >0.5). Hence, density-dependent catchability was applied to area 3r as well. In areas 1r and 4, Mohn´s Rho for recruitment was <0.3. Density-dependent catchability was therefore not applied to these areas. Area 1r, suffered mainly from an SSB retrospective bias that could not be solved in this way. Without going into detail (as area 1r is not part of the inter-benchmark), it should be mentioned that the group agreed that the retro pattern in area 1r was less of a concern than in area 2r. For example, there was no retro-problem last time we saw a large recruitment (2016, same as in area 2r, but in 2r, it was severely overestimated). It also seems that the retro-pattern is improving over time and in the benchmark report, it is noted that the exploitation pattern is estimated-based data only from 2010 and onwards, which may explain the very poor retro values among the oldest of the peels). Area 4 has very little retro, hence, no concern was raised in respect to this area.

Table 1. SSB and recruitment CVs (in the terminal year) for area 2r and 3r, respectively. CVs from assessment back to 2017 are shown, as well as for the 2020 assessment, with and without density-dependent catchability (power model).

Area 2r	2017	2018	2019	2020 (with power model)	2020 (without power model)
SBB (terminal year)	0.26	0.3	0.26	0.18	0.2
Rec (terminal year)	0.61	0.56	0.65	0.57	0.59
Area 3r	2017	2018	2019	2020 (with power model)	2020 (without power model)
SBB (terminal year)	0.42	0.37	0.33	0.3	0.31
Rec (terminal year)	1.98	1.83	1.84	1.81	1.83

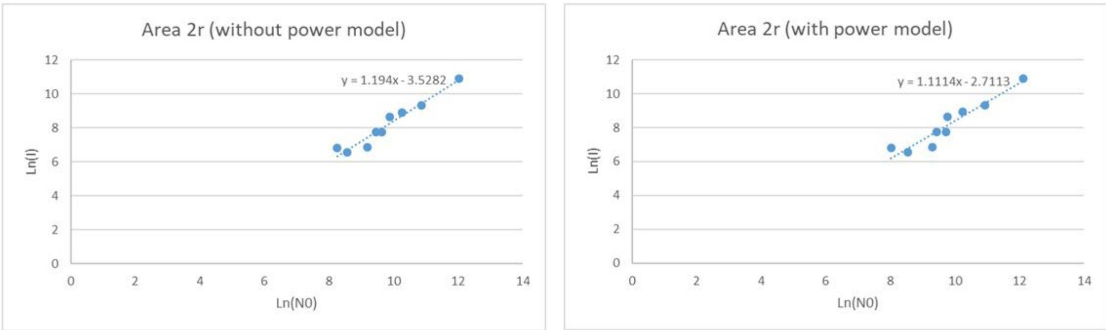


Figure 1. Relationship between Ln(Nage0) and Ln(Indexage0) for area 2r. Since the estimated stock numbers (N) change slightly when implementing density-dependent catchability (power model), the relationship is shown for model runs with and without the power model.

Without power model:

Date: 01/20/20 Start time:15:39:34 run time:1 seconds

objective function (negative log likelihood): 66.5217
Number of parameters: 72
Maximum gradient: 8.44521e-005
Akaike information criterion (AIC): 277.043

With power model:

Date: 01/20/20 Start time:15:43:07 run time:0 seconds

objective function (negative log likelihood): 65.6094
Number of parameters: 73
Maximum gradient: 5.39924e-005
Akaike information criterion (AIC): 277.219

Figure 2. AIC, neg. log likelihood, and maximum gradient for the model with and without density-dependent catchability (power model) for area 2r. Note that the AIC accounts for the number of parameters in the model and therefore should be comparable between models.

Without power model:

sqrt(catch variance) ~ CV:

age	season 1	season 2
0		1.647
1	0.370	0.724
2	0.370	0.724
3	0.800	1.093
4	0.800	1.093

Survey catchability:

	age 0	age 1
Dredge survey 2010-2019	44.180	21.150

sqrt(Survey variance) ~ CV:

	age 0	age 1
Dredge survey 2010-2019	0.41	0.80

With power model:

sqrt(catch variance) ~ CV:

age	season 1	season 2
0		1.662
1	0.374	0.739
2	0.374	0.739
3	0.805	1.090
4	0.805	1.090

Survey catchability:

	age 0	age 1
Dredge survey 2010-2019	1.992	20.938

Stock size dependent catchability (power model)

	age 0	age 1
Dredge survey 2010-2019	1.19	1.00

sqrt(Survey variance) ~ CV:

	age 0	age 1
Dredge survey 2010-2019	0.30	0.79

Figure 3. Catch and survey CVs for the model with and without density-dependent catchability (power model) for area 2r. In the box to the right, there is some additional values and these are the exponent p of the power model (see the equation presented in the main text). All CVs are.

Without power model: With power model:

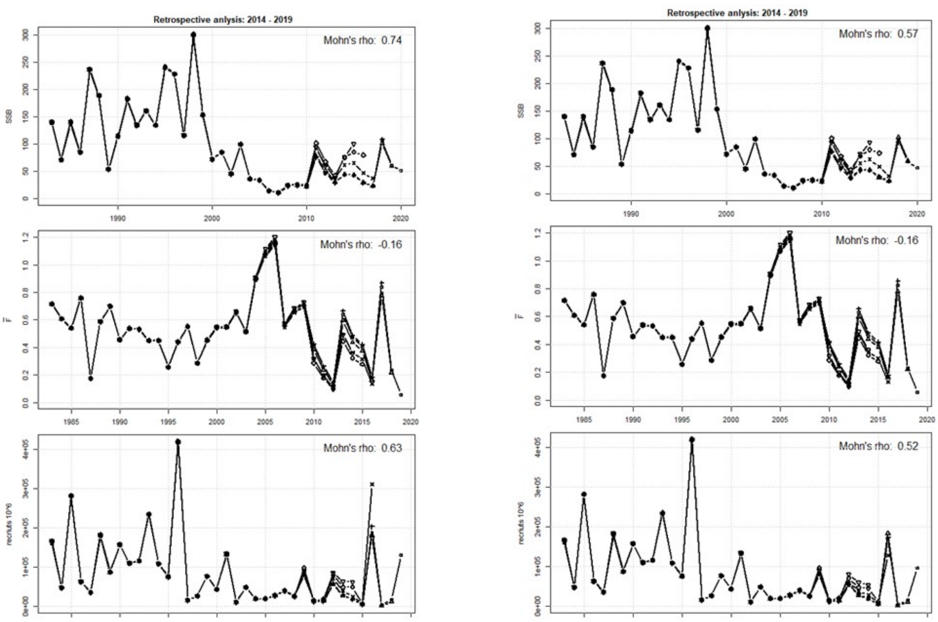


Figure 4. Retrospective bias and Mohn´s Rho for the model with and without density-dependent catchability (power model) for area 2r.

Without power model (area 2r):

```
Recruitment:
i: 1  ( 10878.4 - 16625.3 ) /  16625.3 = -0.345672
i: 2  ( 3431.23 - 3048.06 ) /  3048.06 = 0.1257095
i: 3  ( 311857 - 181505 ) /  181505 = 0.7181731
i: 4  ( 6125.36 - 5175.84 ) /  5175.84 = 0.1834523
i: 5  ( 61453.8 - 17729.8 ) /  17729.8 = 2.46613
rho= 0.6295587   n= 5
```

```
SSB:
i: 1  ( 59.25962 - 60.81911 ) /  60.81911 = -0.02564151
i: 2  ( 108.5742 - 104.4721 ) /  104.4721 = 0.0392646
i: 3  ( 37.21519 - 22.78592 ) /  22.78592 = 0.6332538
i: 4  ( 79.63221 - 29.33589 ) /  29.33589 = 1.714497
i: 5  ( 99.75585 - 42.93954 ) /  42.93954 = 1.32317
rho= 0.7369088   n= 5
```

With power model (area 2r):

```
Recruitment:
i: 1  ( 10382.2 - 15267.7 ) /  15267.7 = -0.3199893
i: 2  ( 5482.65 - 3868.33 ) /  3868.33 = 0.417317
i: 3  ( 127716 - 168519 ) /  168519 = -0.242127
i: 4  ( 10750.8 - 5289.93 ) /  5289.93 = 1.032314
i: 5  ( 52637.6 - 19509 ) /  19509 = 1.698119
rho= 0.5171268   n= 5
```

```
SSB:
i: 1  ( 59.75356 - 58.43638 ) /  58.43638 = 0.02254048
i: 2  ( 93.74566 - 98.14536 ) /  98.14536 = -0.04482843
i: 3  ( 32.09734 - 23.13853 ) /  23.13853 = 0.3871815
i: 4  ( 73.87109 - 31.36566 ) /  31.36566 = 1.355159
i: 5  ( 93.03417 - 43.37529 ) /  43.37529 = 1.144866
rho= 0.5729836   n= 5
```

Figure 5. Values going into the Mohn´s Rho for the model with and without density-dependent catchability (power model) for area 2r. i:1 represents the most recent peel etc. and i:3 is the 2016 recruitment, where the recruitment index was also indicating a large recruitment (just like the present assessment year). Note that after introducing the power model, the recruitment retro in i:3 has disappeared, which was considered an important improvement and an indication that with the power model, we will reduce the risk to overestimate recruitment in 2019 from the survey index (as happened the 2017 assessment year, when the 2016 survey indicated a high recruitment).

Without power model: With power model:

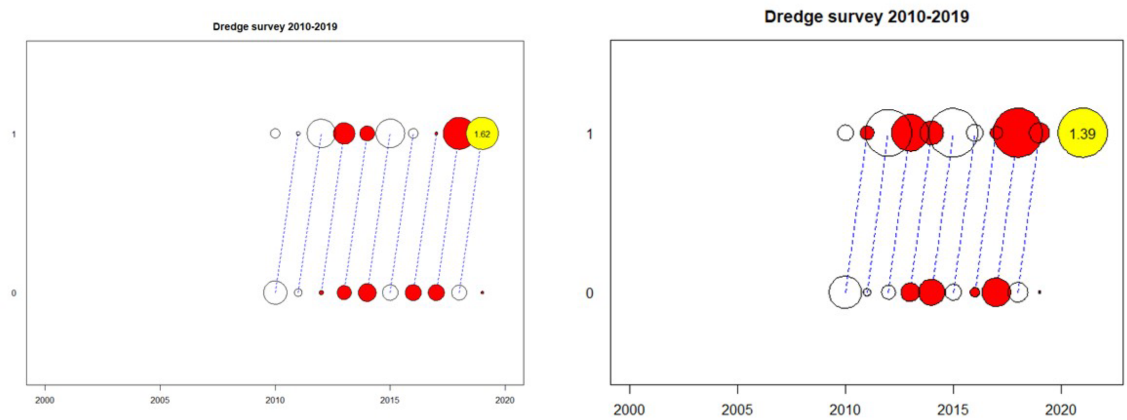


Figure 6. Survey residuals for the model with and without density-dependent catchability (power model) for area 2r. (log(observed CPUE)- log(expected CPUE). "Red" dots show a positive residual. The size of the yellow bubble and the number provides a scale for comparison.

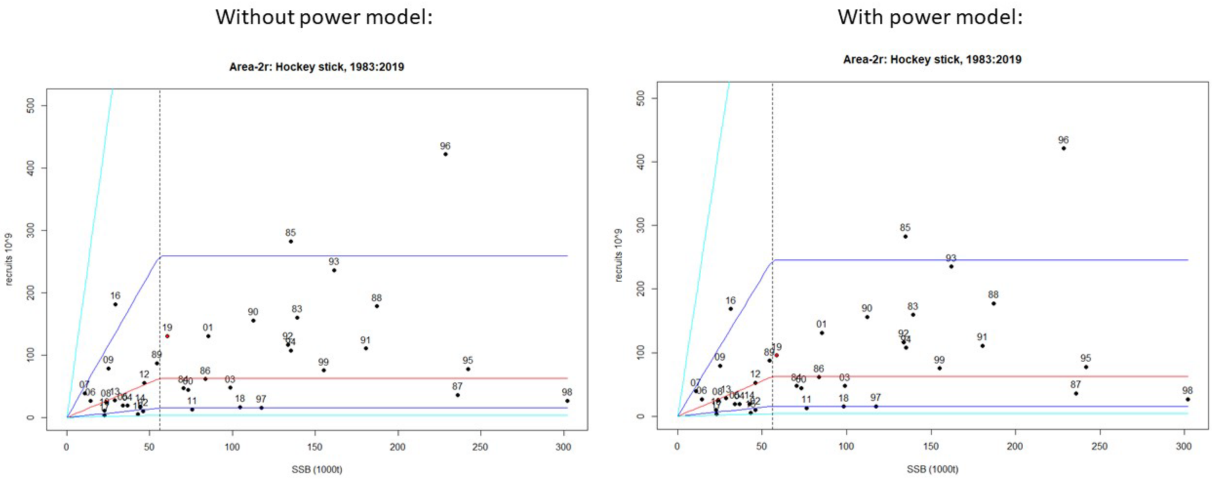


Figure 7. Stock–recruitment plot for the model with and without density-dependent catchability (power model) for area 2r. The vertical line represent B_{lim}.

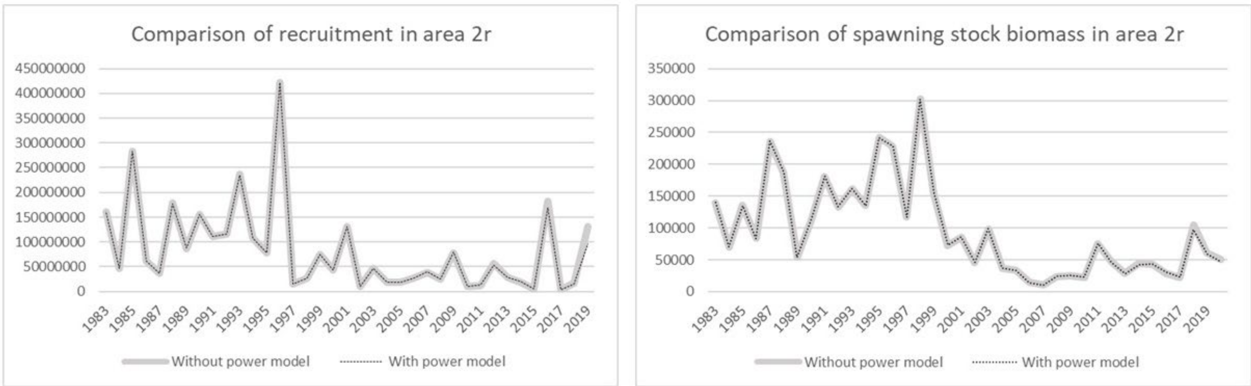


Figure 8. Time-series of recruitment and spawning–stock biomass for the model with and without density-dependent catchability (power model) for area 2r. Note that the survey time-series starts in 2010.

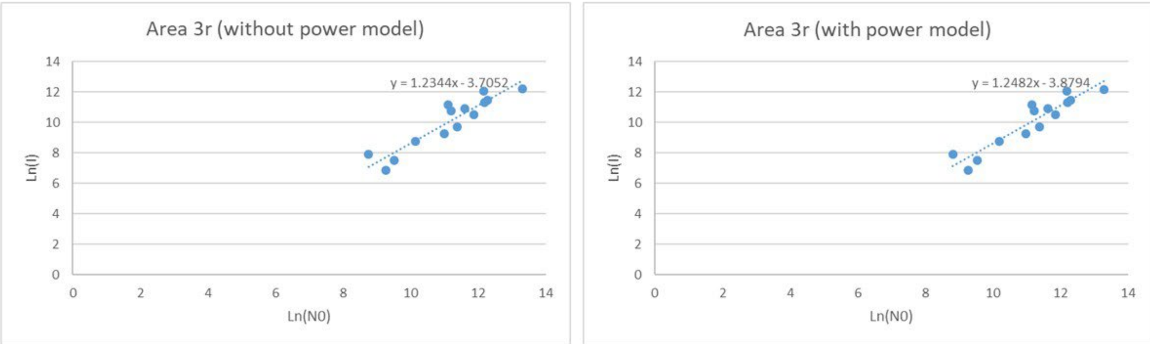


Figure 9. Relationship between Ln(Nage0) and Ln(Indexage0) for area 3r. Since the estimated stock numbers (N) change slightly when implementing density-dependent catchability (power model), the relationship is shown for model runs with and without the power model.

Without power model:

Date: 01/22/20 Start time:16:40:05 run time:2 seconds

objective function (negative log likelihood): 113.271
Number of parameters: 58
Maximum gradient: 7.76889e-005
Akaike information criterion (AIC): 342.543

With power model:

Date: 01/23/20 Start time:11:26:37 run time:0 seconds

objective function (negative log likelihood): 111.639
Number of parameters: 59
Maximum gradient: 8.75414e-005
Akaike information criterion (AIC): 341.279

Figure 10. AIC, neg. log likelihood and maximum gradient for the model with and without density-dependent catchability (power model) for area 3r.

Without power model:

sqrt(catch variance) ~ CV:

season		
age	1	2
0		1.158
1	0.663	1.021
2	0.663	1.021
3	1.054	1.202
4	1.054	1.202

Survey catchability:

	age 0	age 1	age 2	age 3	age 4
Acoustic survey		2.929	5.722	4.656	4.656
Dredge survey 2004-2019	0.751	0.751			

sqrt(Survey variance) ~ CV:

	age 0	age 1	age 2	age 3	age 4
Acoustic survey	0.53	0.53	0.92	0.92	
Dredge survey 2004-2019	0.70	0.88			

With power model:

sqrt(catch variance) ~ CV:

season		
age	1	2
0		1.155
1	0.658	1.027
2	0.658	1.027
3	1.053	1.205
4	1.053	1.205

Survey catchability:

	age 0	age 1	age 2	age 3	age 4
Acoustic survey		2.868	5.602	4.527	4.527
Dredge survey 2004-2019	0.534	0.534			

Stock size dependent catchability (power model)

	age 0	age 1	age 2	age 3	age 4
Acoustic survey		1.00	1.00	1.00	1.00
Dredge survey 2004-2019	1.03	1.00			

sqrt(Survey variance) ~ CV:

	age 0	age 1	age 2	age 3	age 4
Acoustic survey	0.55	0.55	0.92	0.92	
Dredge survey 2004-2019	0.65	0.83			

Figure 11. Catch and survey CVs for the model with and without density-dependent catchability (power model) for area 3r.

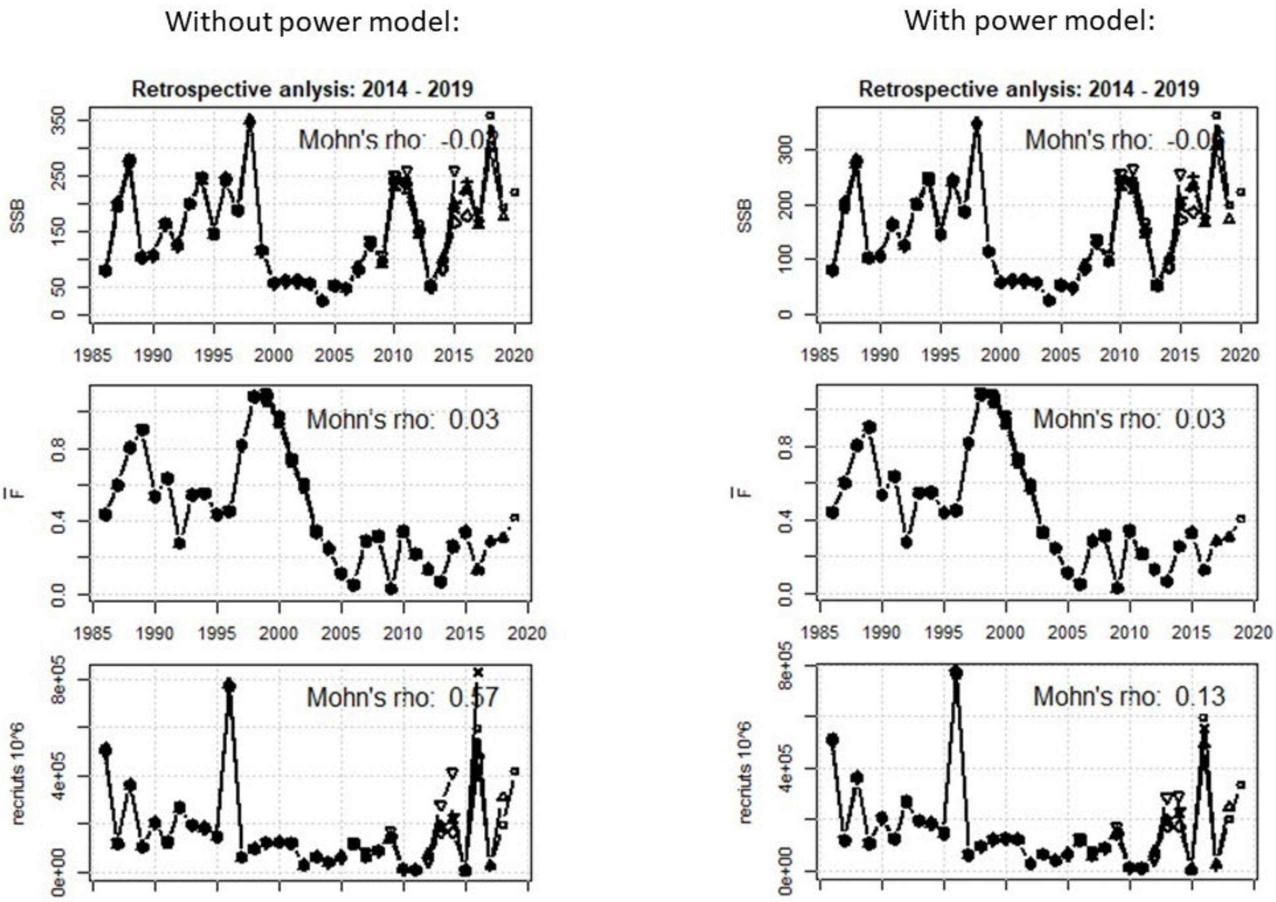


Figure 12. Retrospective bias and Mohn´s Rho for the model with and without density-dependent catchability (power model) for area 3r.

Without power model (area 3r):

```
Recruitment:
i: 1  ( 307897 - 197350 ) /  197350 = 0.5601571
i: 2  ( 29713.3 - 25665.2 ) /  25665.2 = 0.1577272
i: 3  ( 830830 - 596096 ) /  596096 = 0.3937856
i: 4  ( 11413.7 - 6337 ) /  6337 = 0.8011204
i: 5  ( 412973 - 214970 ) /  214970 = 0.9210727
rho= 0.5667726   n= 5
```

```
SSB:
i: 1  ( 176.0285 - 194.7471 ) /  194.7471 = -0.09611748
i: 2  ( 293.9806 - 359.4738 ) /  359.4738 = -0.1821919
i: 3  ( 188.3012 - 171.644 ) /  171.644 = 0.09704542
i: 4  ( 177.5137 - 226.2903 ) /  226.2903 = -0.2155487
i: 5  ( 260.3053 - 198.2617 ) /  198.2617 = 0.3129379
rho= -0.01677496   n= 5
```

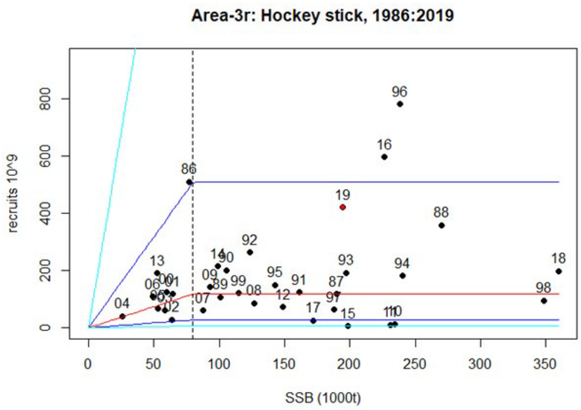
With power model (area 3r):

```
Recruitment:
i: 1  ( 246811 - 202344 ) /  202344 = 0.2197594
i: 2  ( 21670.9 - 26435.5 ) /  26435.5 = -0.1802349
i: 3  ( 556049 - 593191 ) /  593191 = -0.0626139
i: 4  ( 8836.04 - 6697.68 ) /  6697.68 = 0.3192688
i: 5  ( 291367 - 218094 ) /  218094 = 0.3359698
rho= 0.1264298   n= 5
```

```
SSB:
i: 1  ( 173.1032 - 197.4261 ) /  197.4261 = -0.1232
i: 2  ( 301.2808 - 361.8926 ) /  361.8926 = -0.1674855
i: 3  ( 170.0374 - 175.2091 ) /  175.2091 = -0.02951747
i: 4  ( 186.1389 - 232.8677 ) /  232.8677 = -0.2006667
i: 5  ( 253.9661 - 204.2003 ) /  204.2003 = 0.2437107
rho= -0.05543179   n= 5
```

Figure 13. Values going into Mohn´s Rho for the model with and without density-dependent catchability (power model) for area 3r (i:1 represents the most recent peel, etc.).

Without power model:



With power model:

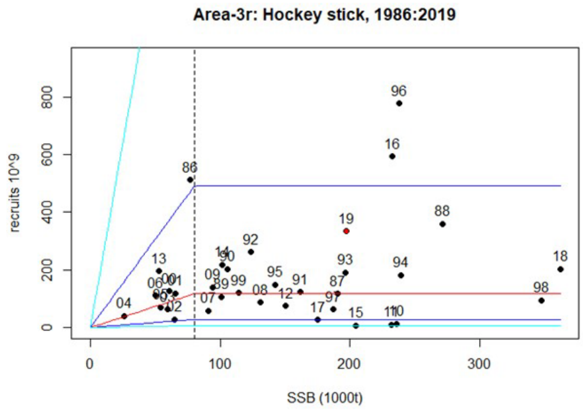


Figure 14. Stock–recruitment plot for the model with and without density-dependent catchability (power model) for area 3r. The vertical line represent B_{lim}.

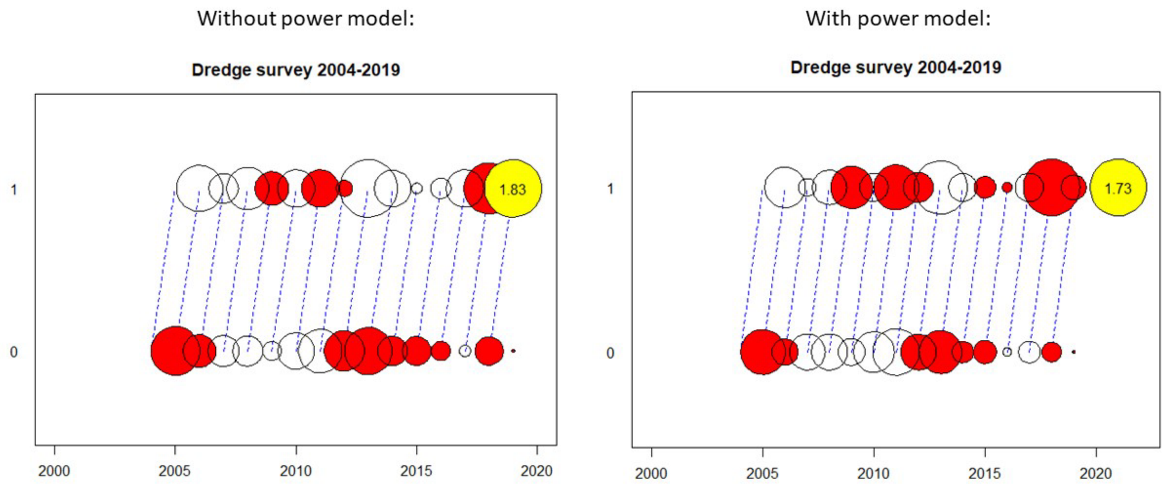


Figure 15. Survey residuals for the model with and without density-dependent catchability (power model) for area 3r. (log(observed CPUE)- log(expected CPUE). "Red" dots show a positive residual. The size of the yellow bubble and the number provides a scale for comparison.

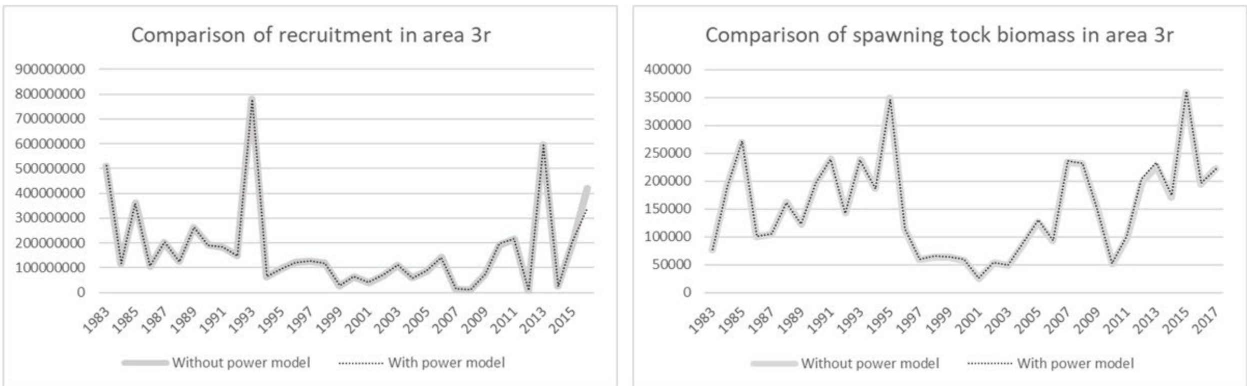


Figure 16. Time-series of recruitment and spawning–stock biomass for the model with and without density-dependent catchability (power model) for area 3r. Note that the survey time-series starts in 2005.

3 Working document Reviews

3.1 Review from Nicola Walker

The working document addresses the problem of a large retrospective pattern in recruitment for the area 2r and 3r sandeel stocks by including a power function representing density-dependence of recruits in the dredge survey. This function has been accepted in the recent benchmark for North Sea sprat on the basis that it is commonly used for short-lived species and improved the retrospective pattern and overall model fit (AIC). Furthermore, there is published evidence of increased survey catchability of sandeel at high density.

3.1.1 Retrospective pattern

According to the draft recommendations from the recent WKFORBIAS 2019 meeting, when a major retrospective pattern exists possible causes and modelling resolutions should be investigated. In the absence of stock-specific thresholds, a retrospective pattern for short-lived species is considered major when Mohn's ϱ is higher than 0.30 or lower than -0.22 and two or three retrospective peels fall outside the confidence bounds of the most recent assessment. The confidence bounds are not presented in Figures 4 and 12, so this part cannot be evaluated.

For the area 2r stock, introducing density-dependent survey catchability improves the retrospective pattern in both SSB and recruitment, but does not reduce Mohn's ϱ to within the threshold. Given that SSB $<B_{lim}$ the decision tree would suggest providing advice over downgrading the assessment or making an adjustment.

Strictly speaking, Mohn's ϱ thresholds apply to SSB so the area 3r stock would not be considered to exhibit a major retrospective pattern (ϱ_{SSB} = -0.02 although this is difficult to see in Figure 12). However, introducing density-dependent survey catchability makes a substantial improvement to the retrospective pattern in recruitment for only a small cost to that of SSB (ϱ_{SSB} = -0.05).

3.1.2 Model fits

For the area 2r stock, introducing density-dependent survey catchability improves the negative log likelihood and survey cvs for a small cost to the AIC and catch cvs (neligible to 0 and 1 d.p respectively), although a large change to the estimated catchability for age 0 is noted.

For the area 3r stock the estimated density-dependent survey catchability is weak (1.03) but it's inclusion clearly improves the overall model fit.

For both stocks, changes to dredge survey residuals would be easier to evaluate if plotted on the same scale between alternate assessment models (Figures 6 and 15). Changes to the stock–recruit relationship, SSB and recruitment time-series appear minimal except for a downscaling of recruitment in the final year; however, peaks in either time-series are somewhat confounded by the different line thicknesses between assessments (Figures 8 and 16).

3.1.3 Conclusions

Including density-dependent survey catchability improves the retrospective pattern for sandeel in area 2r but not does not solve the problem. As the stock is below B_{lim}, I view the more precautionary assessment to outweigh the slight increase in AIC and support the model change.

In terms of SSB, sandeel in area 3r does not exhibit a major retrospective pattern. However, including density-dependent survey catchability improves both the retrospective pattern on recruitment and overall model fit, so I find the model change defensible.

3.2 Review from Pia Schuchert

The document addresses the issue of a strong retrospective pattern in the assessment of sandeels in areas 2r and 3r. The pattern is evident in the SSB and the recruitment; however, this document focusses on the recruitment. There has been published evidence that sandeel recruitment shows density-dependence in dredge surveys. The authors include a power function into the recruitment function to account for this effect.

3.2.1 Retrospective pattern

Major retrospective patterns have been consistently an issue in stock assessment and advice. A decision tree in the new guidelines states that, if there is a Mohn's Rho of greater than 0.2 (for short-lived species as the sandeel greater than 0.3 or lower than -0.22) and at least two of the peels fall outside the confidence interval, modelling resolutions should be considered. In case of the sandeel 2r and 3r assessments Mohn's Rho are >0.5 for SSB and recruitment.

The power function was applied to the recruitment in areas 2r and 3r. Inclusion in area 3r resulted in a considerable improvement the retrospective recruitment pattern, from a Mohn's Rho of 0.567 to 0.126 (Figures 12 and 13). The Mohn's Rho now falls inside the acceptable range, while at the same time, the survey residuals and model fit improved slightly (Figures 11 and 15).

The situation is different in region 2r. The implementation of the power model indeed improves the retrospective pattern of the recruitment and the SSB, however even with the improved model Mohn's Rho for both, SSB and recruitment remains outside the acceptable range (0.57 and 0.52 respectively) (Figures 4 and 5).

The standards set at the WKFORBIAS (2019) meeting then ask to check whether two or more of the last five peels of the assessment are outside the confidence bounds. Without the confidence bounds in Figures 4 and 12, it is difficult to estimate whether this is the case or not. Looking at Figure 5 it might seem that the greatest retrospective patterns stem from earlier rather than the later peels. Regarding the model fit the advanced assessment including the power model has got a slight cost in c.v. and AIC, which is negligible given the improvement in the retrospective pattern.

The changes in SSB, stock–recruitment and F in both areas appear minimal, with only a slight observed decline in the final year recruitment.

3.2.2 Conclusion

The recruitment retrospective pattern in area 3r is considerably improved by the introduction of the power function; indeed all retrospective patterns are now below Mohn's Rho threshold with only minor changes in the model outcomes and slight improvement in model fit. The density-dependence has been seen as an ecological factor in the sandeel population. I therefore support the introduction of the power function into the model.

The situation is slightly different in area 2r. Introduction does improve the retrospective pattern of both, SSB and recruitment. However, the pattern is still considerable larger than Mohn's Rho at 0.57 and 0.52. SSB has been below B_{lim} in 14 of the past 16 years and the introduction of the power function has caused a very slight decline in model fit. The introduction of density-dependent recruitment improves the representation of the ecological recruitment, and while it does not solve the issue of the retrospective pattern considerably improves it. I therefore support the model change.

Annex 1: List of participants

Name	Institute	Country	E-mail
Valerio Bartolino Chair	Institute of Marine Research	Sweden	valerio.bartolino@slu.se
Mikael van Deurs	DTU Aqua	Denmark	mvd@aqua.dtu.dk
Espen Johnsen		Norway	Espen.johnsen@hi.no
Sarah Millar	ICES	Denmark	Sarah-louise.millar@ices.dk
Claus Reedtz Sparrevohn		Denmark	crs@pelagisk.dk
Anna Rindorf	DTU-Aqua	Denmark	ar@aqua.dtu.dk
Pia Schuchert Invited Expert	AFBINI	UK	Pia.Schuchert@afbini.gov.uk
Nicola Walker Invited Expert	Cefas	UK	nicola.walker@cefas.co.uk