IMPACT OF REVISED IBTS SURVEY INDICES ON THE NORWAY POUT ASSESSMENT OUTPUT AND SUSTAINABILITY REFERENCE POINTS (Blim AND Fcap) (AD HOC)

VOLUME 2 | ISSUE 91

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 I

ICES Scientific Reports

Volume 2 | Issue 91

IMPACT OF REVISED IBTS SURVEY INDICES ON THE NORWAY POUT ASSESSMENT OUTPUT AND SUSTAINABILITY REFERENCE POINTS (BLIM AND FCAP) (AD HOC)

Recommended format for purpose of citation:

Brooks, M and Nielsen, J. R. 2020. Impact of revised IBTS survey indices on the Norway pout assessment output and sustainability reference points (Blim and Fcap) (Ad hoc).
ICES Scientific Reports. 2:91. 51 pp. http://doi.org/10.17895/ices.pub.7504

Authors

Mollie Brooks • J. Rasmus Nielsen

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

Contents

1 Introduction

This report presents output from exploratory Norway pout stock (Nop347) assessment runs with special focus on potential change in the MSY and PA sustainability reference points of using the revised IBTS survey indices in DATRAS compared to the previously used indices. Furthermore, the report presents output from exploratory Management Strategy Evaluation (MSE) with consequences for the precautionary F_{cap} of changed biomass reference points.

The report is subdivided into 5 main sections covering presentation of the revised IBTS indices, the results from different exploratory assessment runs in relation to potential change in reference points, and the output from exploratory Management Strategy Evaluation (MSE) of the changed reference points according to a precautionary F_{cap}.

2 Comparison of Previous and Revised Norway pout IBTS Survey Indices

Comparison of old and new IBTS Q1 and Q3 Survey time series with Norway pout indices by age

Table 1 and **Figures 1–2** below show the different IBTS quarter one (Q1) and IBTS quarter three (Q3) survey indices for the Norway pout stock (Nop347) by year and age as extracted from DATRAS.

In general, there is good consistence between the new and old indices, however, it appears that for some ages and years there are extensive differences that in a few cases are close to an order of magnitude. This is especially the case for the older age classes influencing the spawning stock biomass.

Table 1. **The previous (old) and revised (new) IBTS quarter one (Q1) survey indices for Norway pout stock by age and year.**

Year	Quarter	Stock	Version	0	1	2	3	4	5	Version	0	1	2	3	4	5
1972	1	NS_NorPout	old	NA	2578,379	872,136	3,152	NA	NA	new	NA	2588,855	856,062	7,984	NA	NA
1974	1	NS_NorPout	old	NA	25556,97	390,541	24,117	1,291	0,133	new	NA	25559,45	387,954	23,798	1,703	0,15
1975	1	NS_NorPout	old	NA	4573,295	1879,796	4,344	14,214	0,429	new	NA	5066,892	1849,636	36,033	19,891	0,456
1976	1	NS_NorPout	old	NA	4411,127	370,892	2,318	NA	NA	new	NA	4421,661	327,67	35,007	NA	NA
1977	1	NS_NorPout	old	NA	6093,355	273,551	42,216	0,019	NA	new	NA	6121,923	237,721	43,81	3,151	NA
1978	1	NS_NorPout	old	NA	1479,377	574,692	47,013	0,264	NA	new	NA	1480,039	565,267	55,777	0,264	NA
1979	1	NS_NorPout	old	NA	2737,608	316,384	74,948	NA	0,015	new	NA	2737,111	316,157	75,566	NA	0,121
1980	1	NS_NorPout	old	NA	3276,829	550,163	28,843	4,02	0,038	new	NA	3273,839	551,663	30,274	4,073	0,044
1981	1	NS_NorPout	old	NA	1092,244	377,09	14,951	0,178	0,044	new	NA	1091,906	376,835	14,055	0,683	1,073
1982	1	NS_NorPout	old	NA	4537,442	261,84	59,171	1,314	NA	new	NA	4510,766	265,955	80,926	2,147	NA
1983	1	NS_NorPout	old	NA	2258,237	592,073	7,479	3,118	0,056	new	NA	2252,23	592,008	13,338	3,332	0,055
1984	1	NS_NorPout	old	NA	4993,87	981,694	75,477	1,014	0,247	new	NA	4999,685	955,896	89,041	7,203	0,551
1985	1	NS_NorPout	old	NA	2341,565	1429,361	73,158	3,219	0,03	new	NA	2341,555	1400,563	97,798	7,37	0,046
1986	1	NS_NorPout	old	NA	2069,961	382,824	19,957	1,087	0,022	new	NA	2066,071	385,651	18,642	3,452	0,034
1987	1	NS_NorPout	old	NA	3170,71	480,556	61,163	2,895	0,253	new	NA	3170,71	474,613	63,051	6,32	0,883
1988	1	NS_NorPout	old	NA	123,547	721,757	15,021	2,062	0,545	new	NA	123,472	710,207	24,829	3,632	0,793
1989	1	NS_NorPout	old	NA	2018,581	255,343	171,773	2,647	0,227	new	NA	2016,756	253,627	169,705	5,532	2,145
1990	1	NS_NorPout	old	NA	1295,332	747,934	38,611	2,567	NA	new	NA	1295,119	711,9	70,324	7,101	NA
1991	1	NS_NorPout	old	NA	2450,182	712,008	130,183	0,191	0,138	new	NA	2428,442	693,129	156,911	0,961	0,139
1992	1	NS_NorPout	old	NA	5071,202	885,407	31,635	4,288	NA	new	NA	5060,33	860,263	32,907	9,239	NA
1993	1	NS_NorPout	old	NA	2681,504	2644,141	258,496	5,954	7,015	new	NA	2573,999	2643,46	346,171	NA	NA
1994	1	NS_NorPout	old	NA	1838,664	373,575	66,199	2,496	0,223	new	NA	1531,563	372,306	98,661	3,21	0,604
1995	1	NS_NorPout	old	NA	5940,279	784,678	76,765	8,619	0,015	new	NA	5951,374	757,426	85,056	16,47	0,029
1996	1	NS_NorPout	old	NA	922,876	2631,037	227,773	4,698	0,069	new	NA	915,105	2625,861	232,804	12,193	0,478
1997	1	NS_NorPout	old	NA	9699,068	1526,57	669,899	2,861	NA	new	NA	9633,206	1557,297	674,121	33,775	NA
1998	1	NS_NorPout	old	NA	1009,867	5336,449	264,689	86,537	24,849	new	NA	1008,904	5332,124	268,486	87,086	25,791
1999	1	NS_NorPout	old	NA	3526,749	597,123	667,246	3,989	0,286	new	NA	3521,996	600,557	668,239	4,321	0,278
2000	1	NS_NorPout	old	NA	8095,342	1534,851	65	47,568	NA	new	NA	8034,107	1563,22	98,378	47,056	NA
2001	1	NS_NorPout	old	NA	1301,723	2863,245	234,779	3,977	NA	new	NA	1305,987	2805,475	288,165	4,096	NA
2002	1	NS_NorPout	old	NA	1793,285	809,33	879,999	39,757	0,082	new	NA	1784,09	812,305	863,628	61,687	0,742
2003	1	NS_NorPout	old	NA	1239,5	575,349	94,44	88,282	0,484	new	NA	1240,526	573,242	94,384	89,416	0,488
2004	1	NS_NorPout	old	NA	894,263	375,336	33,919	2,25	1,43	new	NA	903,103	364,228	36,566	2,319	1,524
2005	1	NS_NorPout	old	NA	689,606	132,555	37,06	0,702	0,027	new	NA	697,901	123,156	38,332	0,825	0,166
2006	1	NS_NorPout	old	NA	3369,349	141,54	26,366	4,571	1,195	new	NA	3399,972	112,738	22,522	6,901	1,309
2007	1	NS_NorPout	old	NA	1286,127	777,553	22,61	4,863	0,242	new	NA	1287,496	769,233	31,199	3,466	0,439
2008	1	NS_NorPout	old	NA	2352,897	512,037	180,487	9,682	NA	new	NA	2438,209	461,031	153,702	2,874	NA
2009	1	NS_NorPout	old	NA	5479,62	1632,958	151,072	64,301	1,299	new	NA	5553,237	1582,292	122,541	69,663	1,236
2010	1	NS_NorPout	old	NA	4940,787	1466,334	138,273	9,786	1,899	new	NA	4953,748	1438,883	143,349	19,135	1,935
2011	1	NS_NorPout	old	NA	541,069	2251,651	304,029	29,451	3,576	new	NA	544,996	2125,504	347,065	99,93	11,038
2012	1	NS_NorPout	old	NA	997,324	335,995	532,672	28,931	2,41	new	NA	1001,872	326,979	527,058	33,977	7,32
2013	1	NS_NorPout	old	NA	4466,426	519,165	96,563	75,687	0,551	new	NA	4468,54	507,836	101,729	79,459	0,69
2014	1	NS_NorPout	old	NA	812,279	939,435	51,809	18,304	1,5	new	NA	817,566	936,124	47,946	20,218	1,473
2015	1	NS_NorPout	old	NA	6703,722	493,804	140,656	5,498	0,933	new	NA	6637,601	570,081	129,57	7,08	0,281
2016	1	NS_NorPout	old	NA	2416,566	914,528	25,225	4,095	NA	new	NA	2403,51	909,022	41,153	5,009	NA
2017	1	NS_NorPout	old	NA	4357,236	401,126	173,937	3,678	NA	new	NA	4332,483	421,451	173,135	8,909	NA
2018	1	NS_NorPout	old	NA	1157,95	913,65	69,019	3,358	0,469	new	NA	1138,85	850,448	146,701	7,601	0,847
2019	1	NS_NorPout	old	NA	3901,207	294,823	53,725	0,563	0,032	new	NA	3891,543	303,202	54,818	0,767	0,021

Table 2. **The previous (old) and revised (new) IBTS quarter three (Q3) survey indices for Norway pout stock by age and year.**

Year	Quarter	Stock	Version	0	1	2	3	4	Version	0	1	2	3	4
1991	3	NS_NorPout	old	7301,052	1038,696	188,835	1,791	NA	new	7522,989	515,116	485,889	6,379	NA
1992	3	NS_NorPout	old	2559,021	4317,903	633,13	47,676	2,74	new	2559,726	4105,516	740,933	151,071	3,224
1993	3	NS_NorPout	old	4103,932	1831,474	608,477	52,581	3,299	new	4080,262	1506,057	920,861	92,477	0,105
1994	3	NS_NorPout	old	3195,819	704,492	101,593	13,512	0,341	new	3195,645	685,011	113,857	21,146	0,097
1995	3	NS_NorPout	old	2859,57	4440,219	597,372	68,638	1,662	new	2863,622	4106,622	860,133	134,392	2,692
1996	3	NS_NorPout	old	4553,645	763,033	362,416	12,011	0,778	new	4558,905	671,602	419,427	40,532	1,416
1997	3	NS_NorPout	old	489,954	3447,462	235,702	46,499	1,649	new	490,384	3308,438	345,168	76,368	0,919
1998	3	NS_NorPout	old	2931,398	800,941	747,527	12,076	3,012	new	2931,398	791,297	745,224	22,888	4,146
1999	3	NS_NorPout	old	7844,317	2366,567	200,816	94,393	1,463	new	7853,838	2315,676	230,415	106,126	1,47
2000	3	NS_NorPout	old	1643,5	7868,832	281,191	11,258	5,337	new	1643,5	7556,305	589,858	13,618	6,836
2001	3	NS_NorPout	old	2087,776	1273,533	861,513	26,773	2,716	new	2088,628	1163,504	937,889	57,353	4,693
2002	3	NS_NorPout	old	1973,622	766,039	63,634	48,279	0,28	new	1973,622	749,214	76,399	52,324	0,295
2003	3	NS_NorPout	old	1812,158	1063,213	146,247	6,666	4,056	new	1812,256	1014,883	192,957	8,202	3,851
2004	3	NS_NorPout	old	773,252	647,027	152,953	12,321	NA	new	773,342	590,039	208,578	14,261	NA
2005	3	NS_NorPout	old	2678,781	404,227	96,73	16,334	1,307	new	2678,852	395,459	104,033	18,084	1,414
2006	3	NS_NorPout	old	1391,292	1809,196	190,701	12,059	2,609	new	1391,31	1800,018	197,135	14,367	3,585
2007	3	NS_NorPout	old	4150,991	1201,193	447,433	11,446	0,709	new	4151,215	1186,451	429,651	39,588	5,543
2008	3	NS_NorPout	old	3035,265	1643,077	273,771	58,114	1,511	new	3035,278	1610,243	266,909	98,441	1,231
2009	3	NS_NorPout	old	5898,508	2561,831	254,407	11,478	1,204	new	5898,508	2454,251	357,637	13,931	3,094
2010	3	NS_NorPout	old	833,365	4757,28	860,658	22,065	3,365	new	841,893	4780,379	811,813	36,998	5,65
2011	3	NS_NorPout	old	1800,753	473,971	1123,048	59,709	4,084	new	1800,756	474,325	1113,756	64,362	8,375
2012	3	NS_NorPout	old	6416,347	875,134	178,962	130,457	1,67	new	6416,381	828,578	216,975	138,636	2,023
2013	3	NS_NorPout	old	1315,368	2830,775	123,965	12,74	6,255	new	1317,07	2758,836	186,231	17,933	9,527
2014	3	NS_NorPout	old	10238,13	514,197	223,687	7,671	0,598	new	10238,01	479,859	252,583	13,043	0,781
2015	3	NS_NorPout	old	3511,219	4050,917	76,156	19,954	0,523	new	3511,227	3910,592	190,764	46,51	0,653
2016	3	NS_NorPout	old	8965,152	1393,961	277,444	7,655	NA	new	8965,152	1386,48	278,76	13,82	NA
2017	3	NS_NorPout	old	4234,159	2551,163	115,672	20,836	0,086	new	4235,289	2502,113	158,206	25,29	1,017
2018	3	NS_NorPout	old	6114,945	595,161	186,277	5,734	NA	new	6114,945	578,337	201,354	7,481	NA
2019	3	NS_NorPout	old	6460,517	2245,283	100,832	14,095	0,668	new	6464,219	2203,803	133,717	18,991	0,665

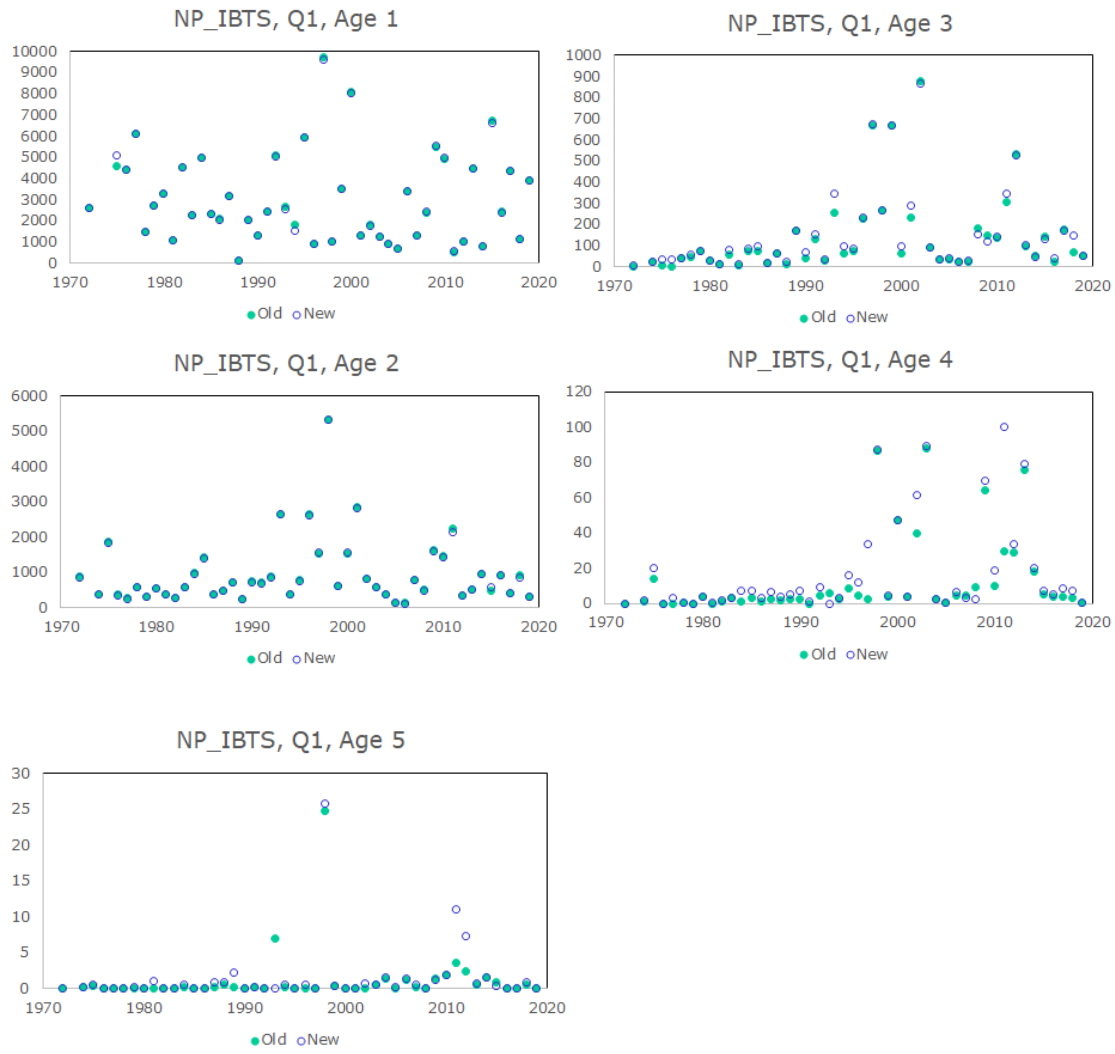

Figure 1. **Comparative plots of new and old IBTS quarter one (Q1) survey indices for Norway pout by age and year. The old indices are indicated with a green filled circle and the new indices are indicated with a blue open circle.**

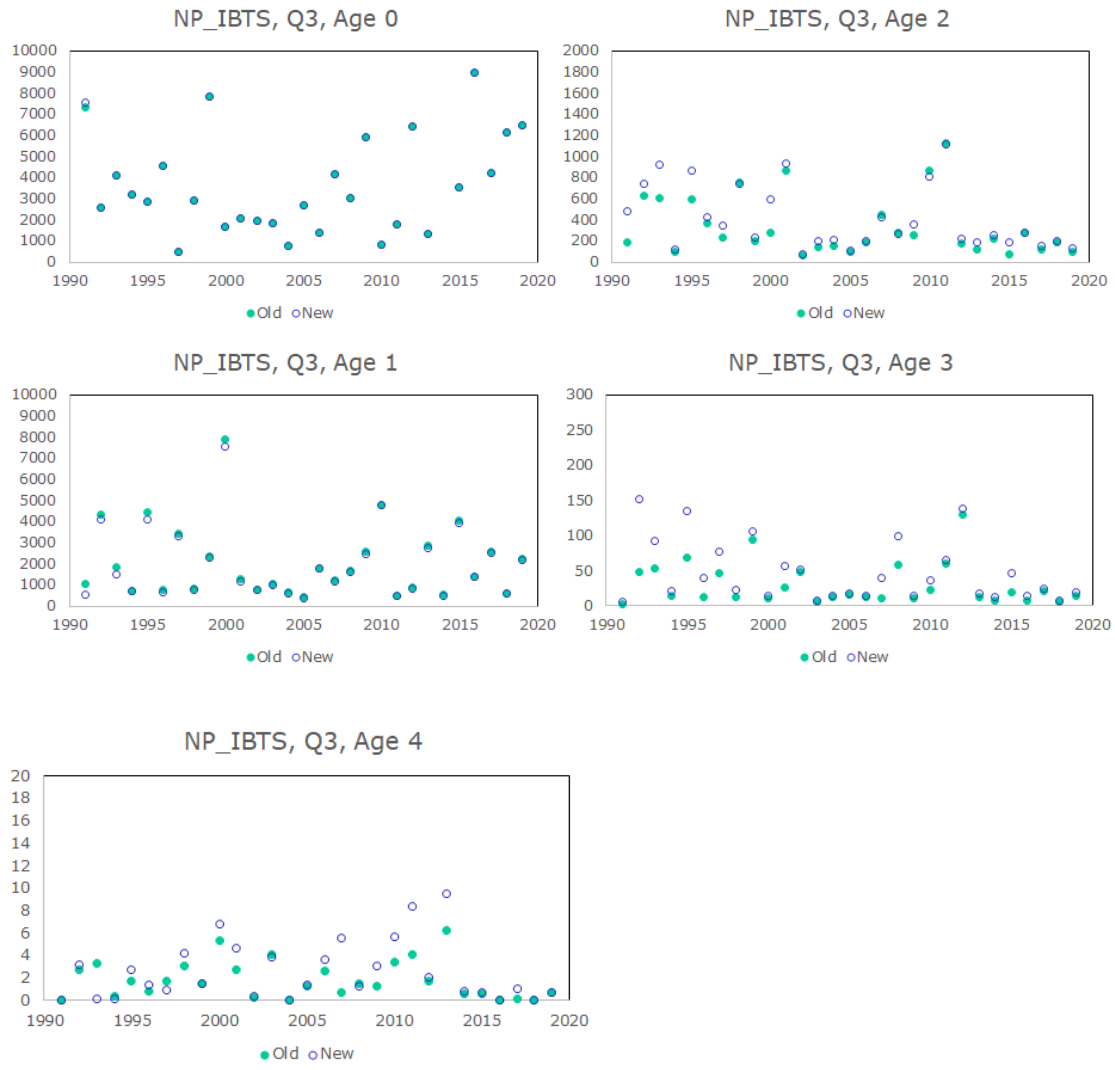

Figure 2. **Comparative plots of new and old IBTS quarter three (Q3) survey indices for Norway pout by age and year. The old indices are indicated with a green filled circle and the new indices are indicated with a blue open circle.**

3 Comparison of Assessment Output from Previous Assessments between the Previous and Revised Norway pout IBTS Survey Indices in relation to Biomass Sustainability Reference Levels

Table 3 below shows the use of the survey indices by age and year in the SESAM assessment.

Table 3 **Norway pout in 4 and 3.aN (Skagerak). Tuning fleets and stock indices and tuning fleets used in the final 2004 benchmark assessment, in the 2005–2015 assessments, as well as in the 2016–2019 assessments based on the 2016 benchmark assessment, compared to the 2003 assessment. (Changes from previous period marked with grey).**

		2003 ASSESSMENT	2004, 2005, April 2006 ASSESSMENT	Sept. 2006 ASSESSMENT	2007-15 ASSESSMENTS	2016-19 ASSESSMENTS
Recruiting season		3rd quarter	2nd quarter (SXSA)	3rd quarter (SMS); 2nd quarter (SXSA	2nd quarter (SXSA), autumn assessm.	3rd quarter SESAM (1984-2019)
Last season in last year		3rd quarter	2nd quarter (SXSA)	3rd quarter (SMS); 2nd quarter (SXSA	2nd quarter (SXSA), autumn assessm.	3rd quarter SESAM (1984-2019)
Plus-group		4+	4+ (SXSA)	None (SMS); 4+ (SXSA)	4+ (SXSA)	3+ (SESAM) (1984-2019)
FLT01: comm Q1						
	Year range	1982-2003	1982-2004	1982-2004	1983-2004, 2006	NOT USED
	Quarter	1	1	1	1	
	Ages	1-3	1-3	1-3	1-3	
FLT01: comm Q2			NOT USED	NOT USED	NOT USED	NOT USED
	Year range	1982-2003				
	Quarter	2				
	Ages	1-3				
FLT01: comm Q3						
	Year range	1982-2003	1982-2004	1982-2004	1983-2004, 2006	NOT USED
	Quarter	3	3	3	3	
	Ages	0-3	1-3	1-3	1-3	
FLT01: comm Q4						
	Year range	1982-2003	1982-2004	1982-2004	1983-2004, 2006	NOT USED
	Quarter	4	4	4	4	
	Ages	0-3	0-3	0-2 (SMS); 0-3 (SXSA)	0-3 (SXSA)	
FLT02: ibtsq1						
	Year range	1982-2003	1982-2006	1982-2006	1983-2015	1984-2019
	Quarter	1	1	1	1	1
	Ages	1-3	1-3	1-3	1-3	1-3
FLT03: egfs						
	Year range	1982-2003	1992-2005	1992-2005	1992-2015	1992-2019
	Quarter	3	Q3 -> Q2	Q3 -> Q2	Q3 -> Q2	3
	Ages	0-3	0-1	0-1	0-1	0-1
FLT04: sgfs						
	Year range	1982-2003	1998-2006	1998-2006	1998-2015	1998-2019
	Quarter	3	Q3 -> Q2	Q3 -> Q2	Q3 -> Q2	3
	Ages	0-3	0-1	0-1	0-1	0-1
FLT05: ibtsq3		NOT USED				
	Year range		1991-2005	1991-2005	1991-2014	1991-2018
	Quarter		3	3	Q3	3
	Ages		2-3	2-3	2-3	2-3

A row of exploratory and comparative assessment were run to compare the assessment output with the previous and revised IBTS indices. This was the most recent September 2019 assessment and the 2017 MSE assessment. Originally it was intended to run the 2016 benchmark assessment, but it was replaced with the 2017 MSE assessment for the below reasons.

The Norway pout 2016 Benchmark assessment included a bug in the code for the SESAM assessment program. This was corrected the first time in the assessment made in 2017 for use in the management strategy evaluations (MSE) on harvest control rules and sustainability of F_{cap} (i.e. in the NP_Sep17_fixC assessment). The fixed C term here means that this bug was corrected and nothing else. Accordingly, instead of running the benchmark assessment program from 2016 then the assessment from 2017 (NP-Sep17_fixC program) was run with the input data used in the benchmark 2016 assessment, but with the new IBTS indices (instead of the original IBTS indices) to check for the change in B_{lim} in relation to the benchmark. This assessment is called NP_Sep17_fixC_Benchmark2016Data_NewIBTS. ‘

The B_{lim} has been shown to vary between yearly assessments (i.e. vary year by year), so this was the decision made during WGNSSK to check sustainability of the currently used B_{lim}. Furthermore, the $B_{lim} = B_{min}$ coming out of this assessment is used in the MSE to evaluate F_{cap} as covered in Section 4.

Table 4 below summarises the resulting $B_{lim} = B_{min}$ at its confidence intervals, i.e. including its uncertainty ranges, from the different exploratory assessments compared with the fixed values obtained from the benchmark 2016 assessment. For all assessments the $B_{lim} = B_{min}$ is consistently the SSB value in the beginning of 4th quarter of 2005, i.e. the 2005.75 value in the respective assessments.

Table 4. Comparison of B_{lim} SSB estimates and their upper and lower confidence limits (C.L) from different assessments with use of previous and revised IBTS time series.

Assessment	Assessment name (www.stockassessment.org)	IBTS Time Series	Year & Quarter	SSB ('000 t)	SSB Upper C.L. ('000 t)	SSB Lower C.L. ('000 t)
Sep 2019 Update assessment	NP_Sep19_b	Previous	2005.75	28279	7117	49441
	NPMar20	New	2005.75	32456	6970	57941
2017 MSE assessment	NP_Sep17_fixC	Previous	2005.75	30742	11233	50252
	NP_Sep17_fixC_NewIBTS	New	2005.75	34999	11013	58985
2017 MSE w. 2016 benchmark data	NP_Sep17_fixC_Benchmark2016Data_NewIBTS	New	2005.75	42573	17046	68100
2016 Benchmark assessment	NorPoutBench2016	Previous	2005.75	39447	17736	61158

It appears from Table 4 that the estimates of B_{lim} is lower for all of the more recent assessments compared to the benchmark assessments. All the assessments with the revised IBTS data estimate B_{lim} within the upper and lower confidence limits of the B_{lim} for the benchmark assessment. Consequently, there is no statistically significant difference between the B_{lim} estimates with previous and new IBTS data time series. In general, it is a uncertain assessment. However, there is a general tendency that the B_{lim} is estimated slightly higher with the new IBTS time series than with the previous IBTS time series. It appears that the benchmark assessment with new IBTS data estimate a $B_{lim} = B_{min}$ at 42 573 tonnes in 2005.75, i.e. a bit higher than the current B_{lim} at 39 447 tonnes from the 2016 Benchmark assessment.

This B_{lim} value of 42 573 tonnes including the new IBTS data time series has been used in the MSE evaluations in Section 4 to check whether the Fcap of 0.7 is still sustainable.

3.1 Exploratory comparison of the 2019 assessment output with old and new survey time series

The results of the compared assessments with respect to biomass reference points are shown in Table 4. The Figure 3 below show comparative plots of the assessment results for spawning stock biomass, total stock biomass, fishing mortality and recruitment when running the most recent SESAM September 2019 update assessment with respectively the previous and the updated time series for the IBTS survey indices.

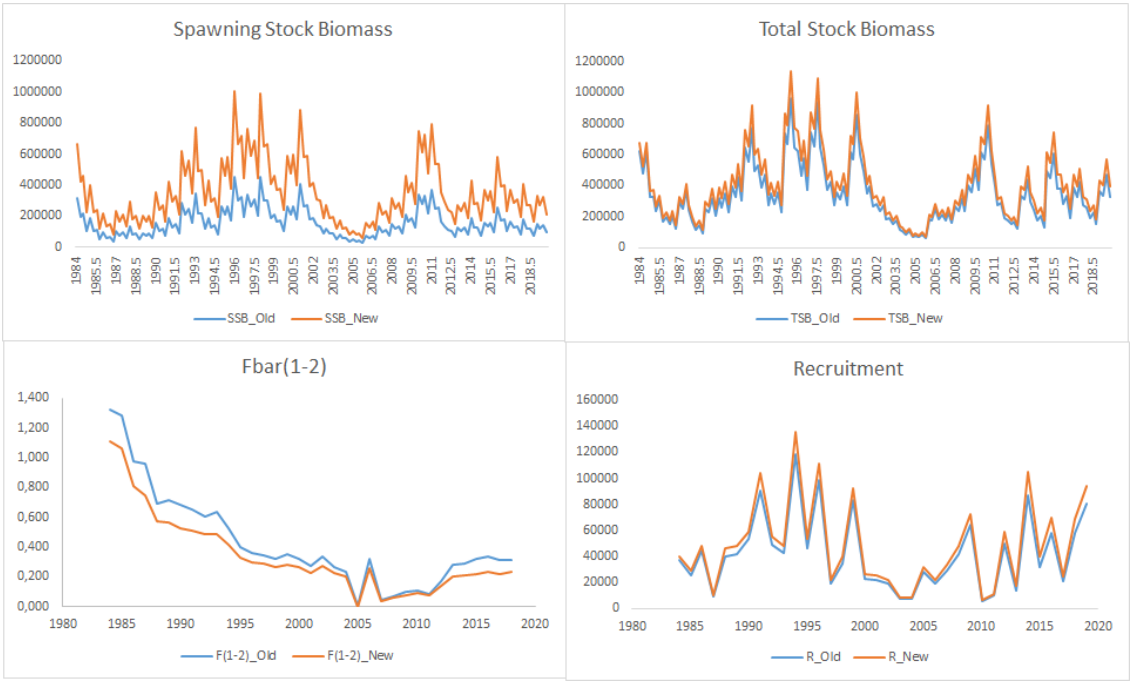

Figure 3. **Comparison of results from the most recent September 2019 assessment with respectively the previous (Old) and revised (New) IBTS indices for spawning stock biomass (SSB), total stock biomass (TSB), fishing mortality ($F_{bar(1-2)}$ and recruitment.**

It appears that the SSB is consistently estimated slightly higher, the F-bar slightly lower, and the recruitment and total spawning stock biomass at approximately the same level over time in the assessment using the new, revised IBTS indices compared to the assessment using the previous IBTS data time series. The changes are all within the uncertainty ranges of the estimates, i.e. there are no significant differences in the assessment results, as appear from Appendix A.1. The perception of the stock and stock development, the performance of the assessment and the stock recruitment relationship is approximately identical between the assessments using the new IBTS data time series compared to the assessment using the previous data time series as also apparent from Appendix A.1.

3.2 Exploratory comparison of 2017 assessment output for MSE with old and new survey time series

A 2017 assessment with corrected catch calculations was used to condition (i.e. to parameterize) the MSE conducted for the 2018 benchmark. Results of the compared assessments with respect to biomass reference points are shown in Table 4. Figure 4 below shows comparative plots of the assessment results for spawning stock biomass, total stock biomass, fishing mortality and recruitment when running the SESAM September 2017 assessment with respectively the previous and the updated time series for the IBTS survey indices.

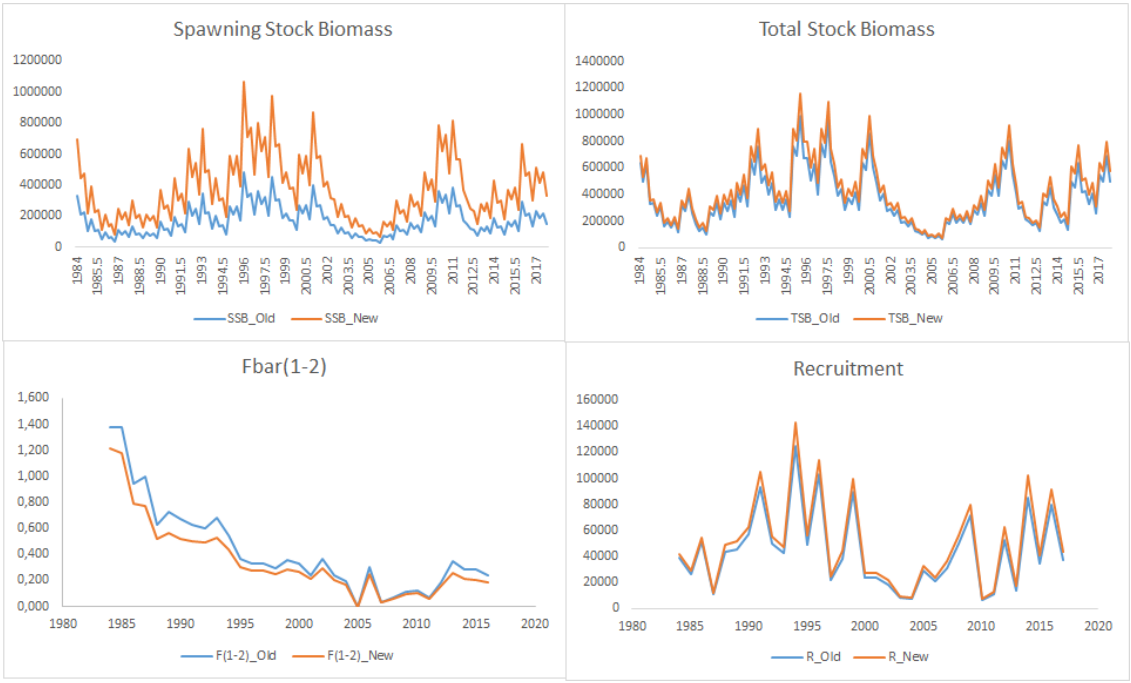

Figure 4. **Comparison of results from the September 2017 assessment with respectively the previous (Old) and revised (New) IBTS indices for spawning stock biomass (SSB), total stock biomass (TSB), fishing mortality ($F_{bar(1–2)}$ and recruitment.**

It appears that the results are very much consistent and with the same general perception of the stock as for the comparison with the September 2019 update assessment as described under section a. above. See Appendix A.2 for further details of the comparison of the September 2017 assessment results.

3.3 Exploratory comparison of the 2016/2017 Benchmark assessment output with old and new survey time series

Results of the compared assessments with respect to biomass reference points are shown in Table 4. Figure 5 below shows comparative plots of the assessment results for spawning stock biomass, fishing mortality and recruitment when running the SESAM 2016/2017 Benchmark assessment with respectively the previous and the updated time series for the IBTS survey indices. The total stock biomass was not calculated in the 2016 Benchmark assessment. Accordingly, there is not shown a comparative plot for total spawning stock biomass here.

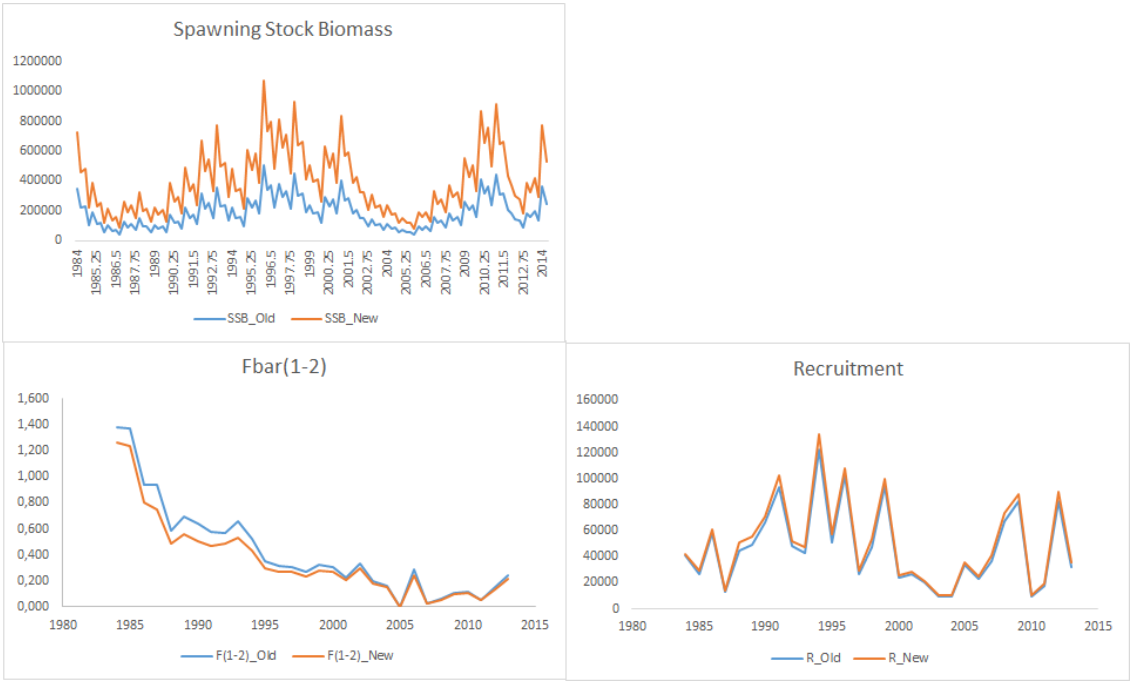

Figure 5. **Comparison of results from the 2016/2017 benchmark assessment with respectively the previous (Old) and revised (New) IBTS indices for spawning stock biomass (SSB), total stock biomass (TSB), fishing mortality ($F_{bar(1–2)}$ and recruitment.**

Again it appears that the results of the comparisons of use of previous and new IBTS data are very much consistent with the comparison using the September 2019 update assessment as described under section a. above. See Appendix A.3 for further details of the comparison of the 2016–2017 Benchmark assessment results.

4 Management Strategy Evaluation of Sustainable Fcap levels

Evaluation of revised Fcap reference points will demand a full scale MSE similar to the one conducted in 2018 and presented in the WKNPOUT report (ICES, 2018) with similar considerations as provided in this report. This will demand extensive resources and will need a full benchmark process to evaluate robustness of reference points according to B_{lim} and Fcap and the necessary considerations among other in relation to assumptions presented herein. Furthermore, refer to recommendations in previous benchmarks. Here we performed a scaled-down MSE of a smaller number of HCRs compared to the 2018 benchmark, with fewer performance statistics, and fewer sensitivity tests.

Each HCR tested contained a maximum TAC (i.e. TAC_{max} in tonnes) and a minimum TAC (i.e. TAC_{min} in tonnes) as well as an Fcap. For the main MSE, we tested HCRs with combinations of TAC_{min} equal to 0, 20k, 30k, or 40k tonnes; TAC_{max} equal to 150k or 200k tonnes, and Fcap equal to 0.3, 0.4, 0.5, 0.6, 0.7, or 0.8. The chosen TAC_{max} of 200k tonnes was approximately the highest catch recorded in the stock's recent history (over the last 25 years). For each HCR, we ran 1000 simulation trials (i.e. replicates).

4.1 MSE Conditioning (i.e. Parameterization)

The MSE followed the same assumptions as in the 2018 benchmark MSE (ICES, 2018) but with values extracted from the new assessment from 2017 run with the new IBTS data as described above in Section 3.2 of this report.

We assumed that recruitment followed a hockey stick model with an inflection point at the lowest estimated SSB in quarter 1 in the new assessment from 2017 run with the new IBTS data as described above in Section 3.2 of this report (and Appendix A.2). SSB estimated in Q1 in 2005 was the lowest at 62 683 tonnes. Therefore, recruitment in the operating model of the MSE is impaired if SSB in Q1 is below 62 683 tonnes. We assumed that for SSB in Q1 above the breakpoint, then recruitment is a random sample from the estimated recruitment including uncertainty from previous years. See ICES, 2018 for details.

We assumed that F_{bar} would not exceed the maximum estimated in the last 20 years of the assessment. This is a form of implementation error as described in ICES, 2018. It assumes that the fleet will not exert more effort than they have in recent history. We let Fhistorical equal to 0.67, which is the maximum 97.5% percentile of F_{bar} estimated by SESAM in the last 20 years according to the 2017 assessment run with new IBTS data (the only assessment used to condition this MSE – see section 3b and Appendix A.2 of current report). This value is the upper confidence interval of the 2013 and 2002 F_{bar}. This assumption had an impact in up to 20% of simulated total realized catches TRCs (Figure A.4.1) and therefore we did a sensitivity test. TRC is TAC with implementation error. To test the sensitivity of risk to this assumption, we separately ran simulations that allowed F_{bar} to go up to 1 with TAC_{min} = 0, TAC_{max} = 200k or 400k tonnes, and Fcap in a range from 0.3 to 0.7.

4.2 Performance Statistics

We calculated a subset of the performance statistics from the 2018 benchmark (Table 5). As described above, a B_{lim} value of 42 573 tonnes SSB in Q4 was used to calculate risk because Q4 is when the escapement strategy targets SSB.

Table 5. Performance statistic descriptions

F_{bar}.median	Median true $F_{age\ 1–2}$
F_{bar}.mean	Mean true $F_{age\ 1–2}$
SSB.median	Median SSB in quarter 4 in tonnes
SSB.mean	Mean SSB in quarter 4 in tonnes
risk3.short.Q4	Probability of SSB in quarter 4 is below B_{lim}. The maximum risk in one of the years 2018–2022 is used (ICES Risk type 3)
risk1.long.Q4	Probability of SSB in quarter 4 is below B_{lim}. The average risk in the years 2023–2037 is used (ICES Risk type 1)
risk3.long.Q4	Probability of SSB in quarter 4 is below B_{lim}. The maximum risk in any of the years 2023–2037 is used (ICES Risk type 3)
risk1.long.Q1	Probability of SSB in quarter 1 is below the inflection point in the Hockey-stick SR applied. The average risk in the years 2023–2037 is used (ICES Risk type 1)
atFhist	Probability that the TAC will require a true $F_{age\ 1–2}$ higher than $F_{historical}$ to be taken.
TAC.median	Median TAC in tonnes
TAC.mean	Mean TAC in tonnes
TRC.median	Median Total Realized Catch weight in tonnes (catch taken with a true F capped at $F_{historical}$)

4.3 Results

Our main result is that we found that Fcap = 0.7, TAC_{max} = 200k tonnes, and TAC_{min} = 0 tonnes is precautionary (Figure 6, Figure 7). Sensitivity tests showed that this result was robust to removing the assumption about implementation error limiting F_{bar} to be below Fhistorical and increasing TAC_{max} to 400k tonnes (Figure 8, Figure 9).

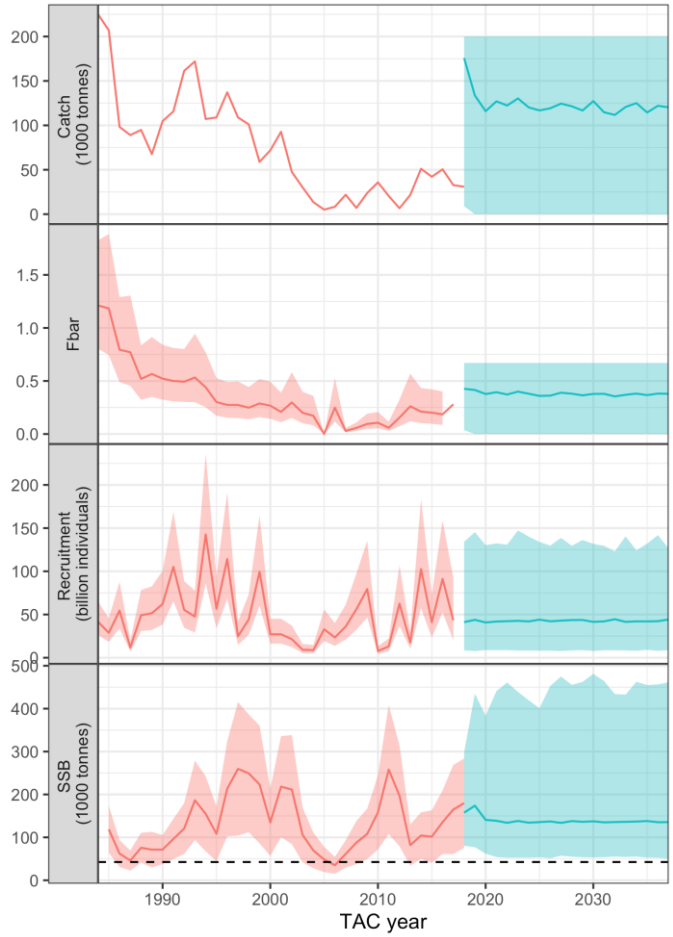

Figure 6. **Summary result from the SESAM assessment of Norway pout (in red) and scenario values using the escapement strategy with Fcap = 0.7 and TAC_{max} = 200k (in green). Catch is catch weight by TAC year, F_{bar} is the average of quarterly F_{age1-2} within a TAC year, Recruitment is stock number at age zero in the beginning of quarter 3, and SSB is SSB in the beginning of quarter 4. The lines show the median value and the shaded area the 5th and 95th percentiles. The horizontal dashed black line is B_{lim}.**

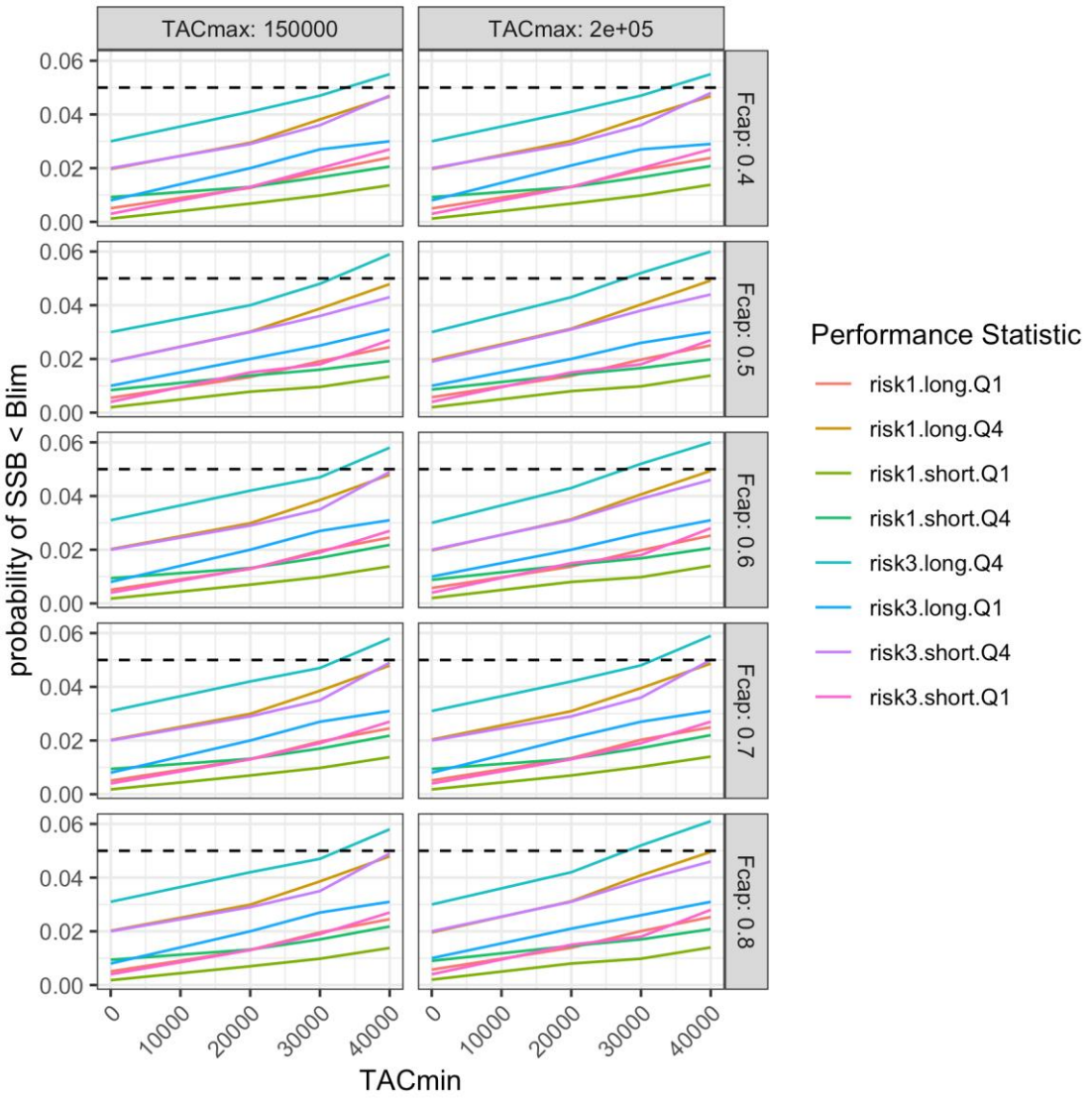

Figure 7. **Risk performance statistics from the MSE with a range of HCRs varying the maximum TAC (columns: TAC_{max}), minimum TAC (x-axis: TAC_{min}), and Fcap (rows). See Table 5 for performance statistic descriptions.**

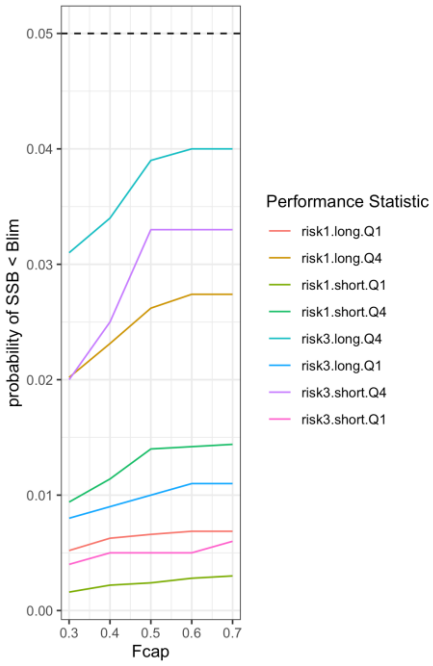

Figure 8. **Sensitivity results. For a range of values of Fcap (x-axis) with TAC_{max} = 400k tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to 1.0. See Table 5 for performance statistic descriptions.**

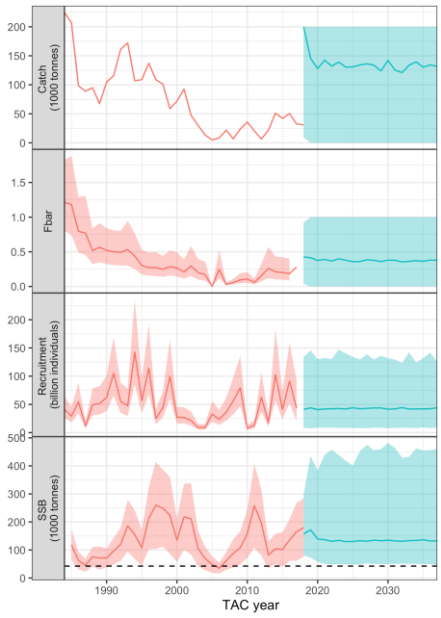

Figure 9. **Summary plot of sensitivity result for Fcap = 0.7, TAC_{max} = 200k tonnes and TAC_{min} = 0 tonnes, where the MSE was rerun while allowing F_{bar} to go up to 1.0. The plot includes the SESAM assessment of Norway pout (in red) and scenario values using the escapement strategy with Fcap = 0.7 and TAC_{max} = 200k (in green). Catch is catch weight by TAC year, F_{bar} is the average of quarterly $F_{age1–2}$ within a TAC year, Recruitment is stock number at age zero in the beginning of quarter 3, and SSB is SSB in the beginning of quarter 4. The lines show the median value and the shaded area the 5th and 95th percentiles. The horizontal dashed black line is B_{lim}.**

Additional performance statistics defined in Table 5 are presented in Appendix A.4.

5 Impact of revised IBTS survey indices on the Norway pout assessment output and sustainability reference points (B_{lim} and F_{cap}) – Part II

5.1 Introduction

This section is a section added after a review of the document "Impact of revised IBTS survey indices on the Norway pout assessment output and sustainability reference points (B_{lim} and F_{cap})". The purpose of this section is to present additional output from exploratory Management Strategy Evaluations (MSE). It is not meant to be read as a stand-alone document and should only be considered in combination with the main parts of this report.

5.2 MSE sensitivity test with infinite TAC_{max} and F_{max} = 0.67 the historical maximum estimated

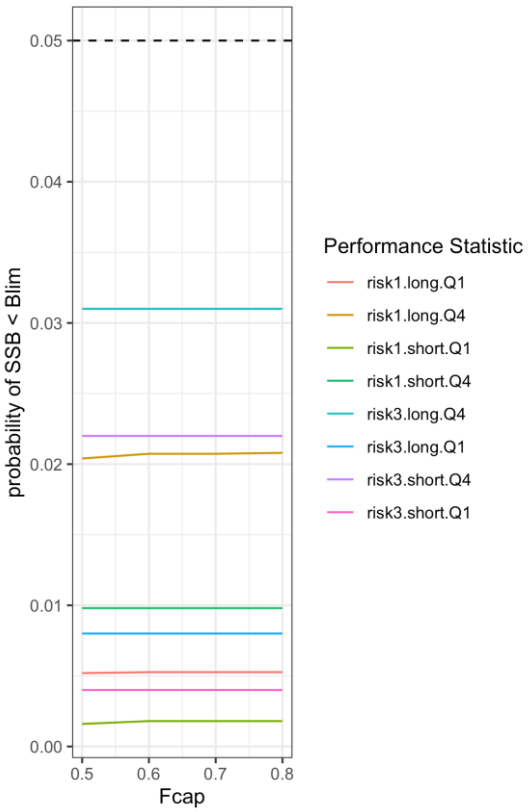

Figure 2.1 **For a range of values of F_{cap} (x-axis) with TAC_{max} = infinity tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to its historical maximum. See Section 4.2, Table 5 for performance statistic descriptions.**

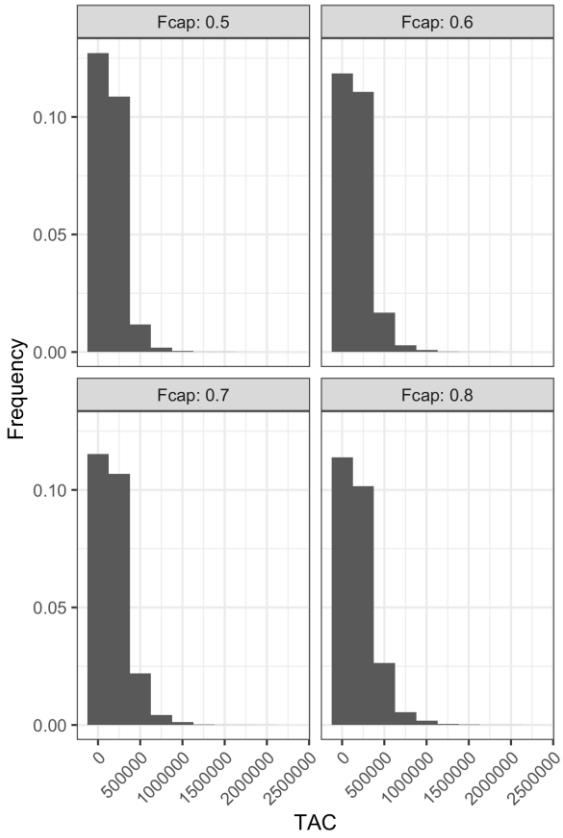

Figure 2.2 **Frequency distribution across replicates and years 2023 to 2037 of TAC depending on F_{cap} with with TAC_{max} = infinity tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to its historical maximum.**

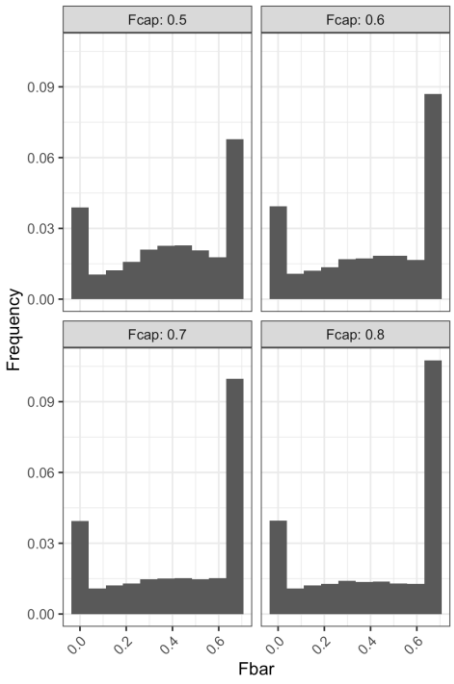

Figure 2.3 **Frequency distribution across replicates and years 2023 to 2037 of F_{bar} depending on F_{cap} with with TAC_{max} = infinity tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to its historical maximum.**

5.3 MSE sensitivity test with infinite TAC_{max} and F_{max} = 1.0

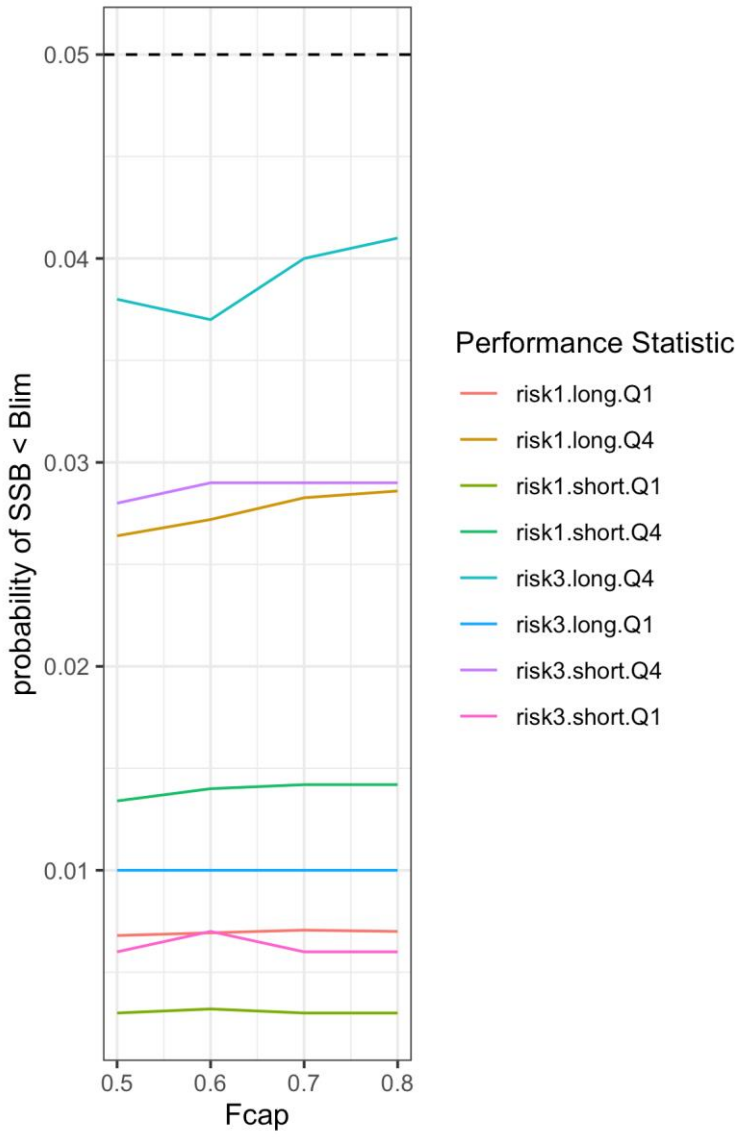

Figure 3.1 **For a range of values of F_{cap} (x-axis) with TAC_{max} = infinity tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to 1.0. See Section 4.2, Table 5 for performance statistic descriptions.**

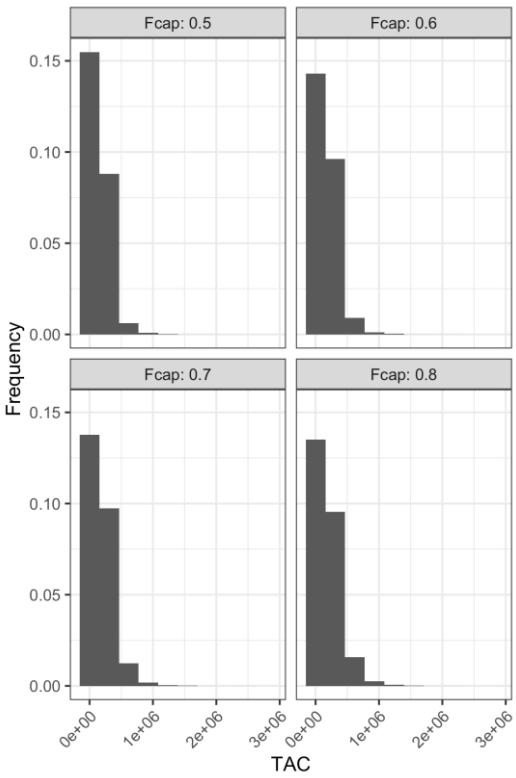

Figure 3.2 **Frequency distribution across replicates and years 2023 to 2037 of TAC depending on F_{cap} with with TAC_{max} = infinity tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to 1.0.**

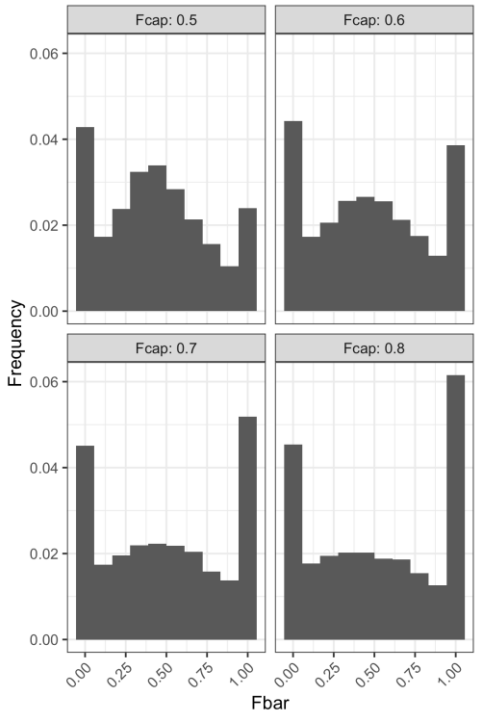

Figure 3.3 **Frequency distribution across replicates and years 2023 to 2037 of F_{bar} depending on F_{cap} with with TAC_{max} = infinity tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to 1.0.**

5.4 MSE sensitivity test with infinite TAC_{max} and F_{max} = 2.0

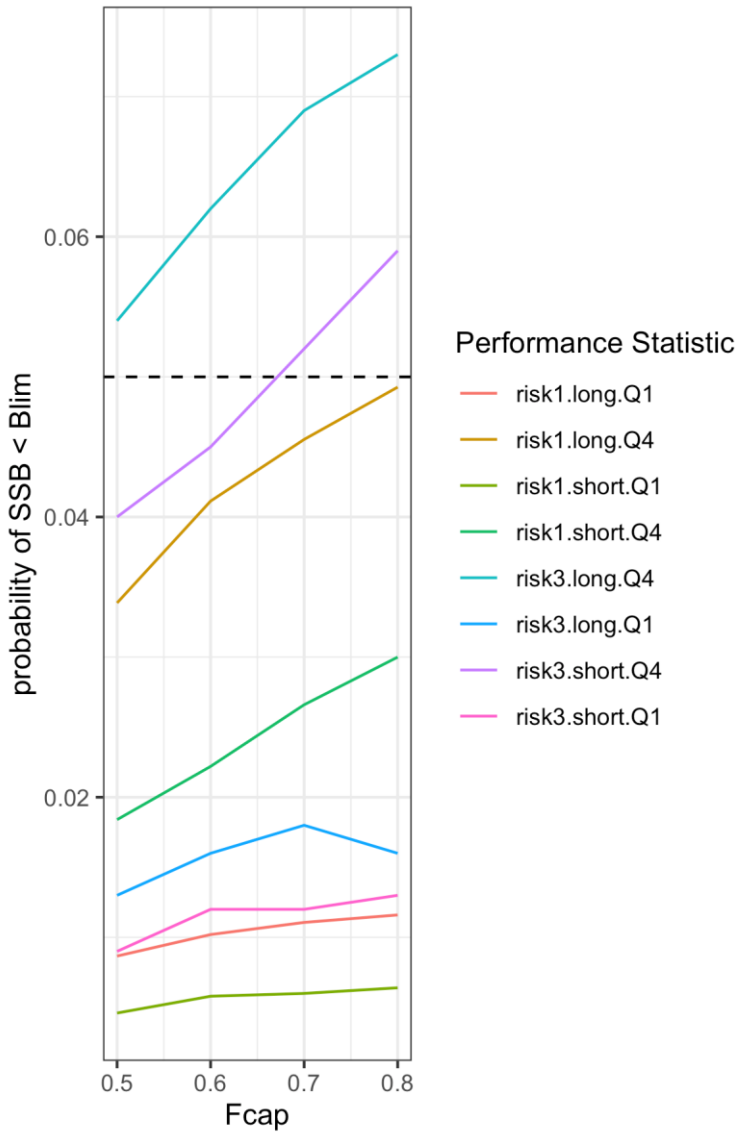

Figure 4.1. **For a range of values of F_{cap} (x-axis) with TAC_{max} = infinity tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to 2.0. See Section 4.2, Table 5 for performance statistic descriptions.**

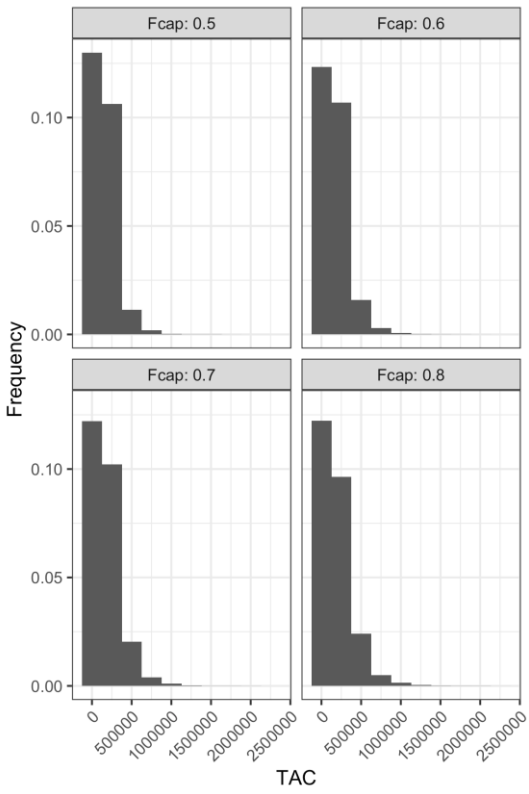

Figure 4.2 **Frequency distribution across replicates and years 2023 to 2037 of TAC depending on F_{cap} with with TAC_{max} = infinity tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to 2.0.**

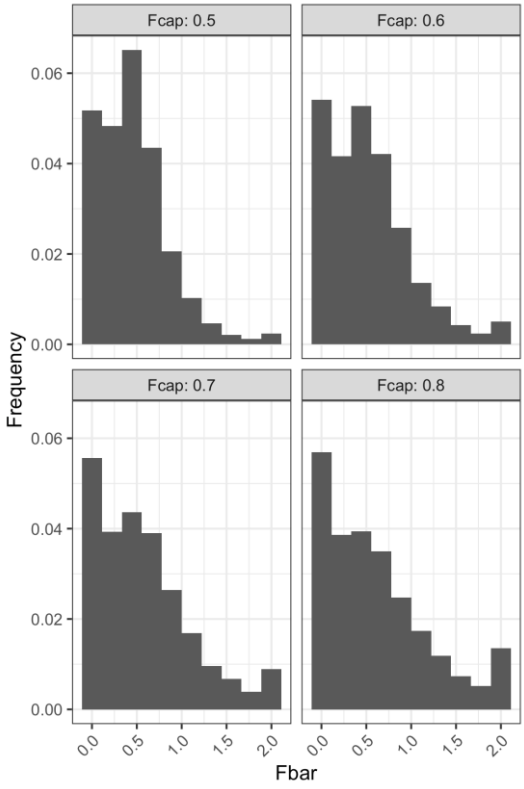

Figure 4.3 **Frequency distribution across replicates and years 2023 to 2037 of F_{bar} depending on F_{cap} with with TAC_{max} = infinity tonnes and TAC_{min} = 0 tonnes, the MSE was rerun while allowing F_{bar} to go up to 2.0.**

5.5 Conclusion

With no limits on TAC, then the assumption of a maximum implementable F has a stronger effect on the simulated stock dynamics. When the maximum implementable F is near F_{cap} (Section 5.2), then F_{cap} has very little effect on the stock dynamics. If we assume that the maximum implementable F is extremely large (2.0 which is more than double the maximum estimated value), then the effect of F_{cap} can be seen again (Section 5.4). With maximum implementable F at either its maximum historical estimate or at 1.0, then all risk statistics still show F_{cap} = 0.7 to be precautionary (figures 2.1 and 3.1 of this section). Furthermore, even with the unrealistically high maximum implementable F, then the only risk that goes above 0.05 (when rounded to the nearest 0.01 units) is risk3.long.Q4 for a F_{cap} = 0,7. The type 3 risk statistics may require more replicates to converge to the true value expected from infinite replicates (WKGMSE2 - ICES, 2019); if needed, this could be investigated in a benchmark. However, the overall result is that risk 1 statistics all indicate precautionarity even under extreme assumptions for high fishing effort.

6 References

ICES. 2018. Report of the Workshop for management strategy evaluation for Norway Pout (WKNPOUT). ICES WKNPOUT Report 2018 26-28 February 2018. ICES HQ, Copenhagen, Denmark. ICES CM 2018 / ACOM:38 Ref: WGNSSK. 96 pp.

ICES. 2019. Workshop on Guidelines for Management Strategy Evaluations (WKGMSE2). ICES Scientific Reports. 1:33. 162 pp. http://doi.org/10.17895/ices.pub.5331

Appendix A.1.

Detailed assessment results from NP_Sep2019 Update Assessment with previous and revised IBTS data time series

Norway Pout Sep. 2019 Update Assessment: Comparative assessment results between NP_Sep19b with previous IBTS data (left hand side plots) and NPMar20 with new IBTS data (right hand side plots). Summary of results – quarterly with uncertainties.

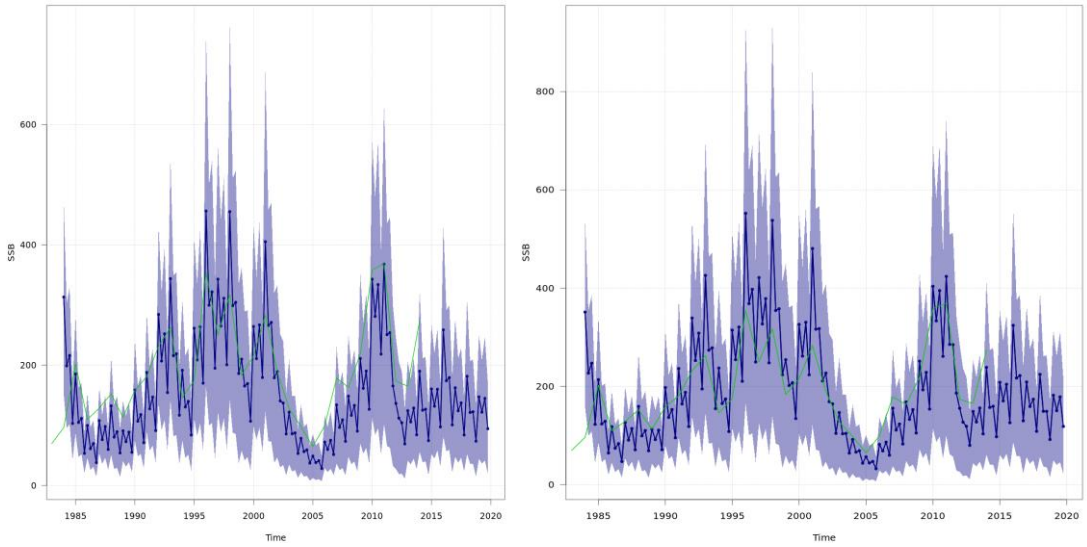

Figure A.1.1 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: SSB (t), quarterly. SESAM baseline run September 2019. Quarterly estimated SSB and confidence interval from SESAM (blue) and SXSA (green, quarter 1 only – connecting lines are interpolations).**
Old left; New right

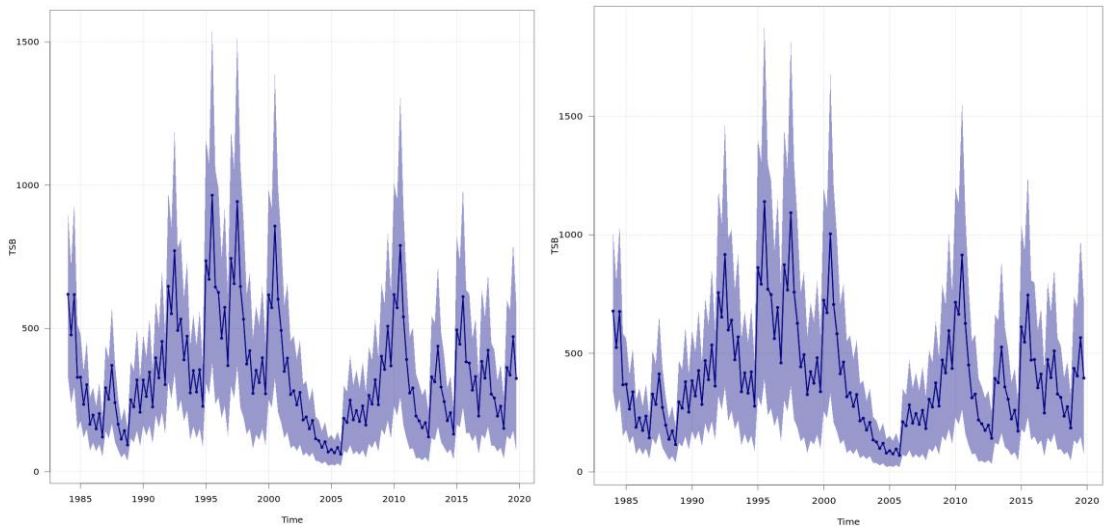

Figure A.1.2 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: TSB (t), quarterly. SESAM baseline run September 2019.**
Old left; New right.

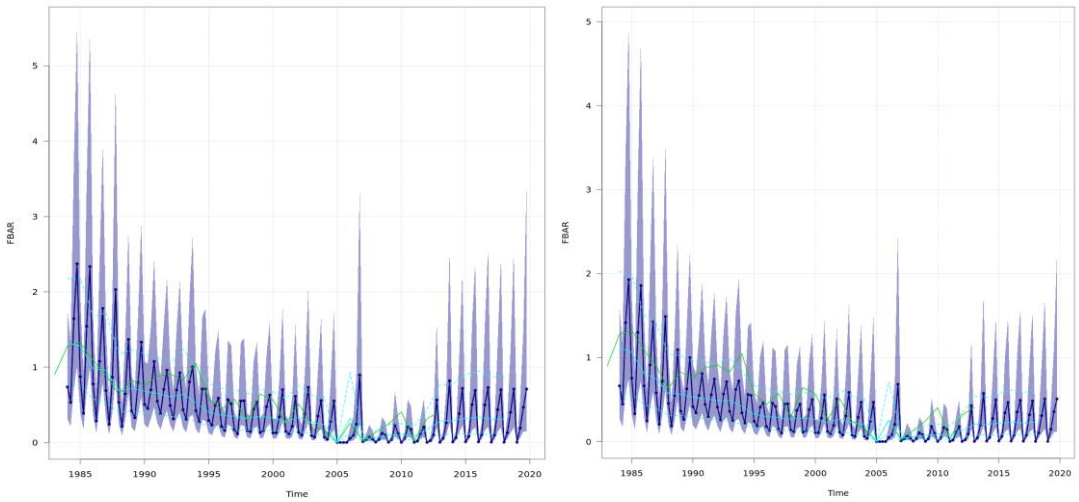

Figure A.1.3 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: $F_{1-2} = F_{bar}$, quarterly. SESAM baseline run September 2019. Blue is quarterly values from SESAM, cyan is the yearly average from SESAM, green is yearly average from SXSA.**
Old left; New right.

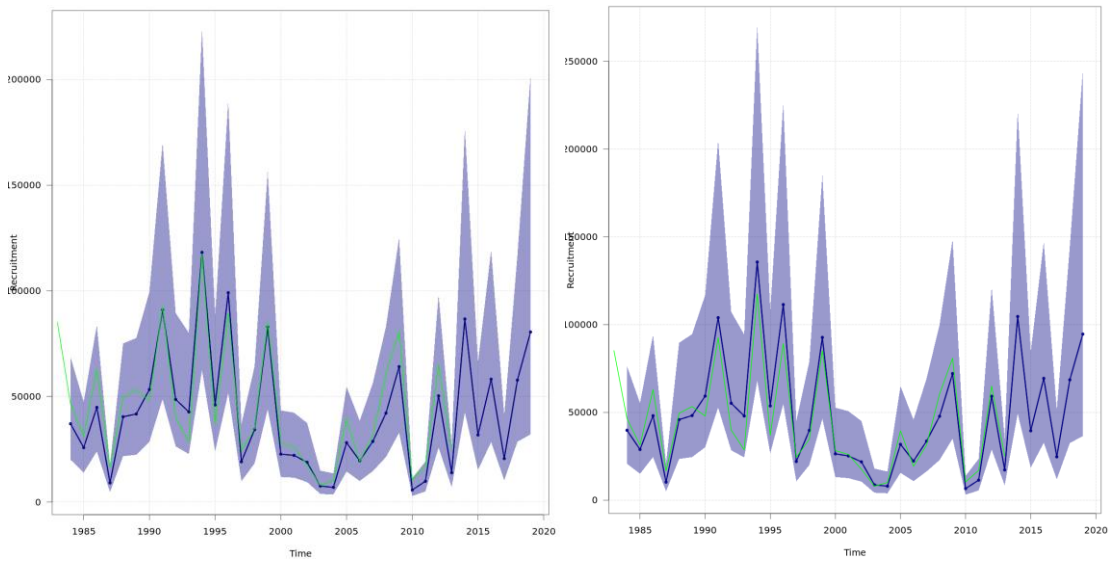

Figure A.1.4 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: Recruitment (millions), yearly. SESAM baseline run September 2019. Blue is SESAM, green is SXSA.**
Old left; New right.

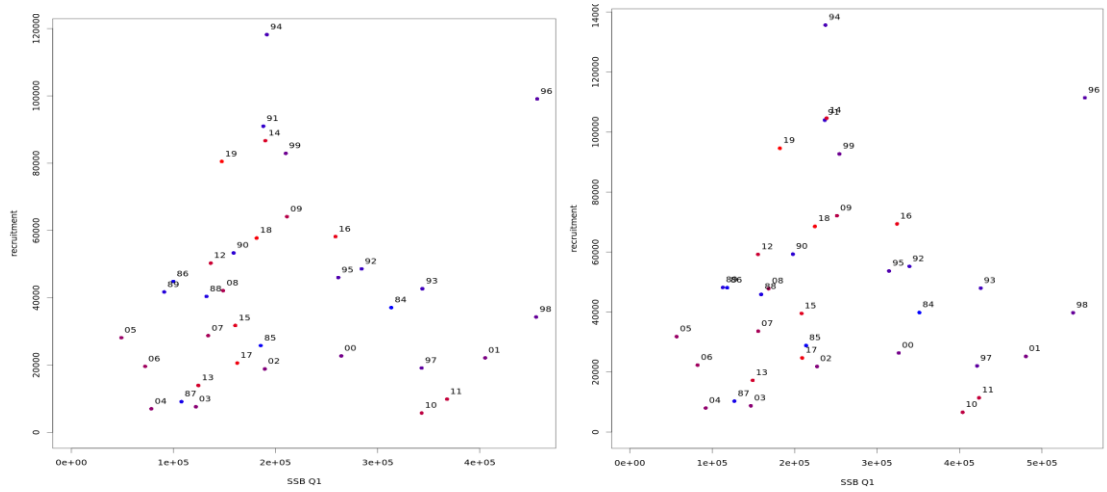

Figure A.1.5 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: Stock (SSB) – Recruitment Plot Quarter 1. SESAM baseline run September 2019.**
Old left; New right.

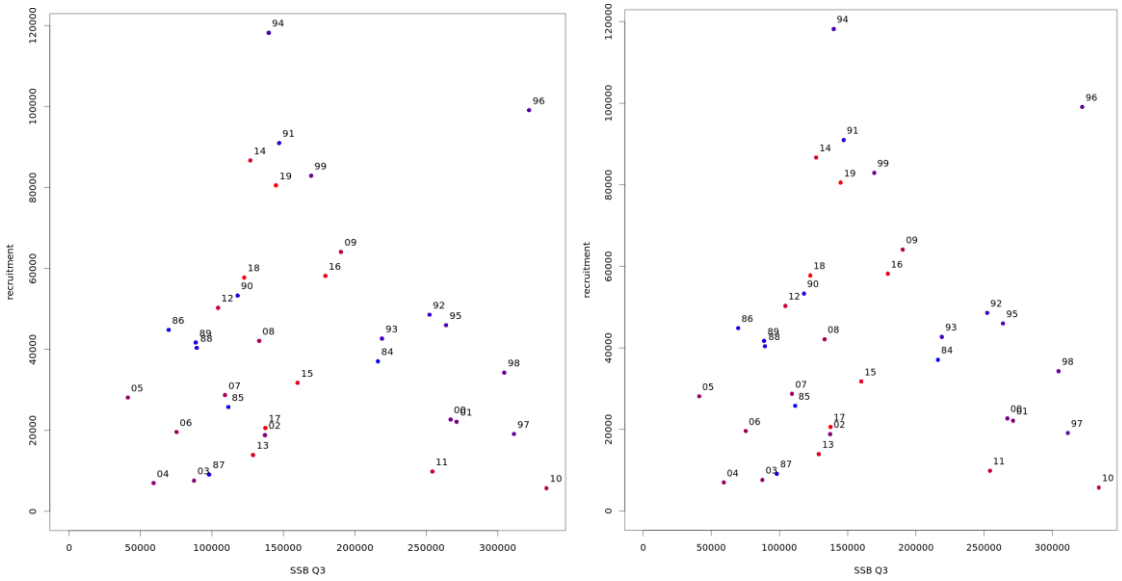

Figure A.1.6 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: Stock (SSB) – Recruitment Plot Quarter 3. SESAM baseline run September 2019.**
Old left; New right.

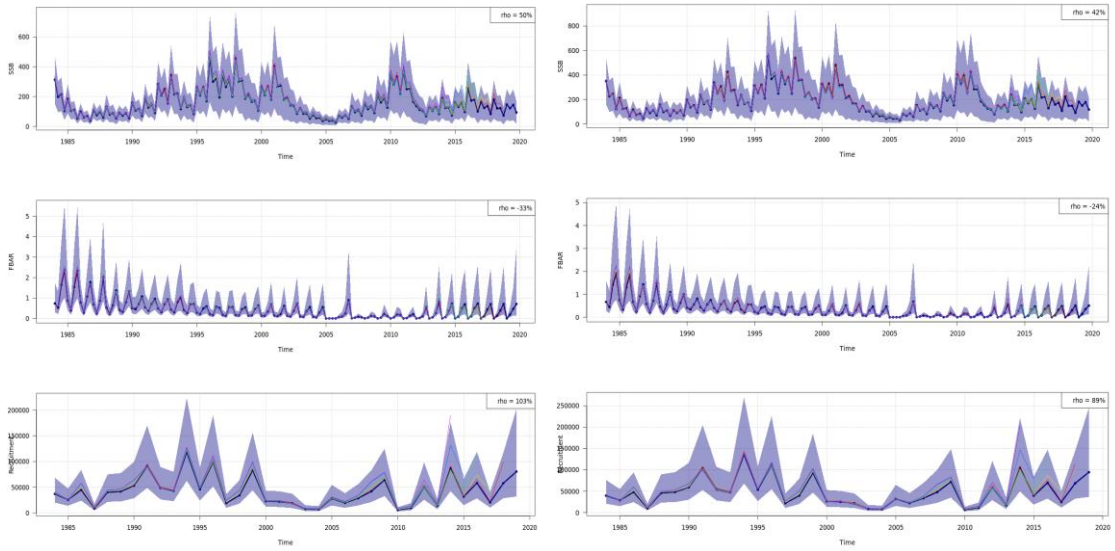

Figure A.1.7 **Norway pout in 4 and 3.aN (Skagerrak). Retrospective plots of baseline SESAM assessment September 2019, with terminal assessment year ranging from 2005–2019. Updated with Mohns Rho.**
Old left; New right.

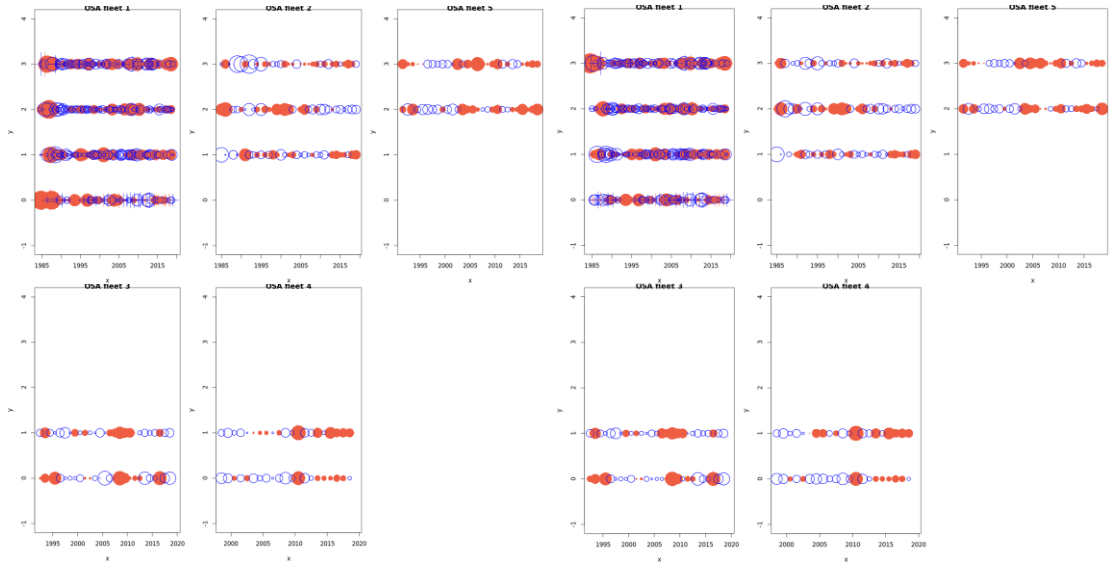

Figure A.1.8 **Norway Pout in 4 and 3.aN (Skagerrak). Assessment Diagnostics Plots by fleet and age: One step ahead residuals (see Berg and Nielsen 2016). SESAM baseline run September 2019. Old left; New right.**

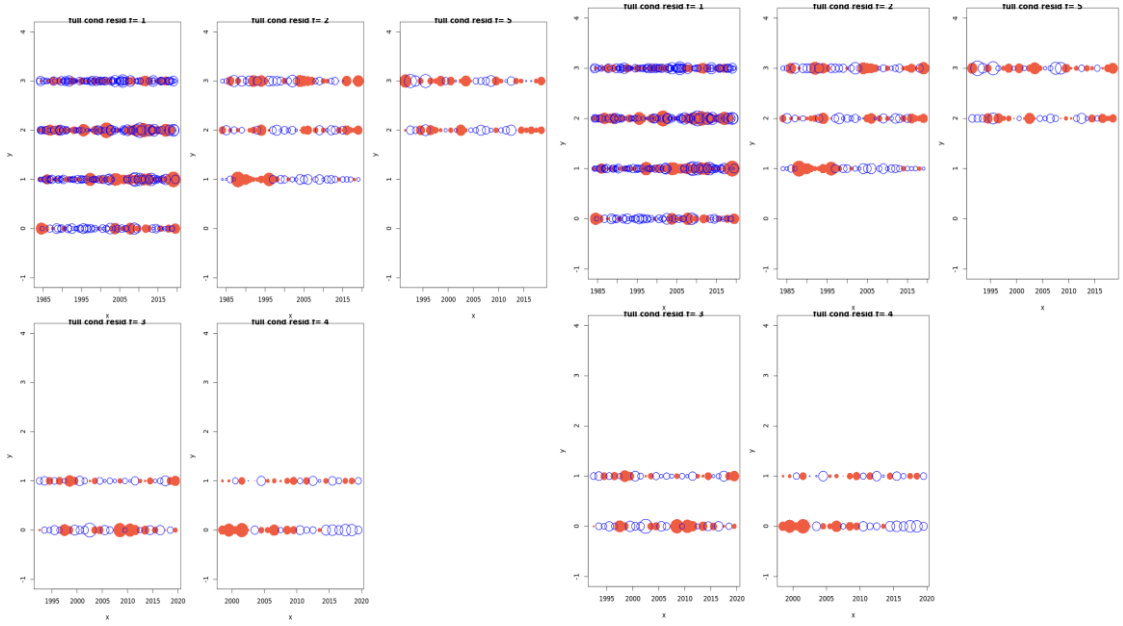

Figure A.1.9 **Norway Pout in 4 and 3.aN (Skagerrak). Assessment Diagnostics Plots: Full conditional residuals or auxiliary residuals by fleet and age (see Berg and Nielsen, 2016). SESAM baseline run September 2019. Old left; New right.**

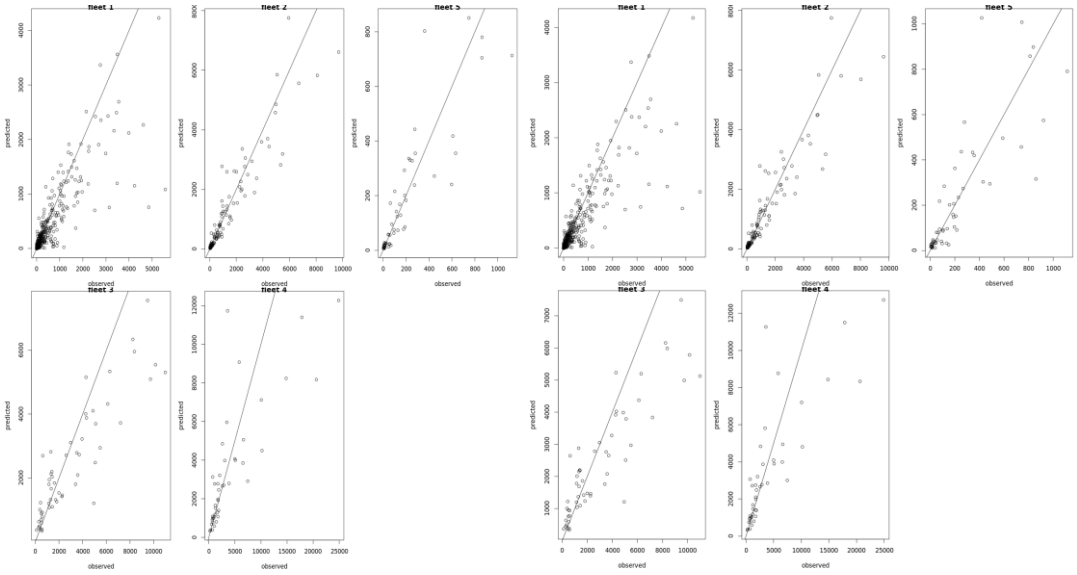

Figure A.1.10 **Norway Pout in 4 and 3.aN (Skagerrak). Assessment Diagnostics Predicted vs Observed catches. Plots by fleet. SESAM baseline run September 2019.**
Old left; New right.

Appendix A.2.

Detailed assessment results from the 2017 MSE Assessment with previous and revised IBTS data time series

Norway Pout 2017 MSE Assessment: Comparative assessment results between NP_Sep17_fixC with previous IBTS data (left hand side plots) and NP_Sep17_fixC_NewIBTS with new IBTS data (right hand side plots). Summary of results – quarterly with uncertainties.

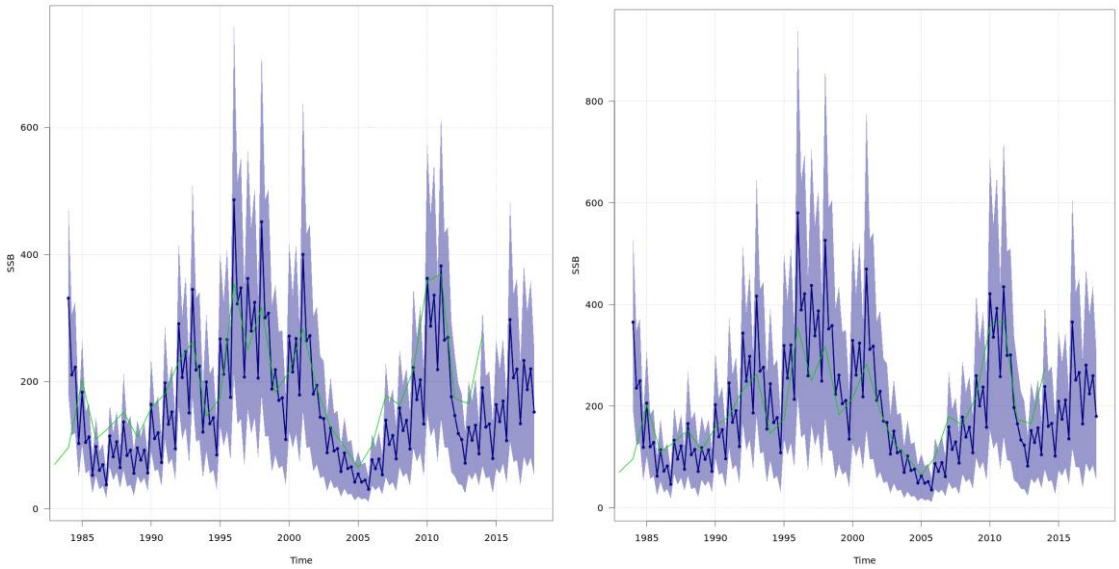

Figure A.2.1 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: SSB (t), quarterly. SESAM 2017 MSE Assessment. Quarterly estimated SSB and confidence interval from SESAM (blue) and SXSA (green, quarter 1 only – connecting lines are interpolations).**
Old left; New right.

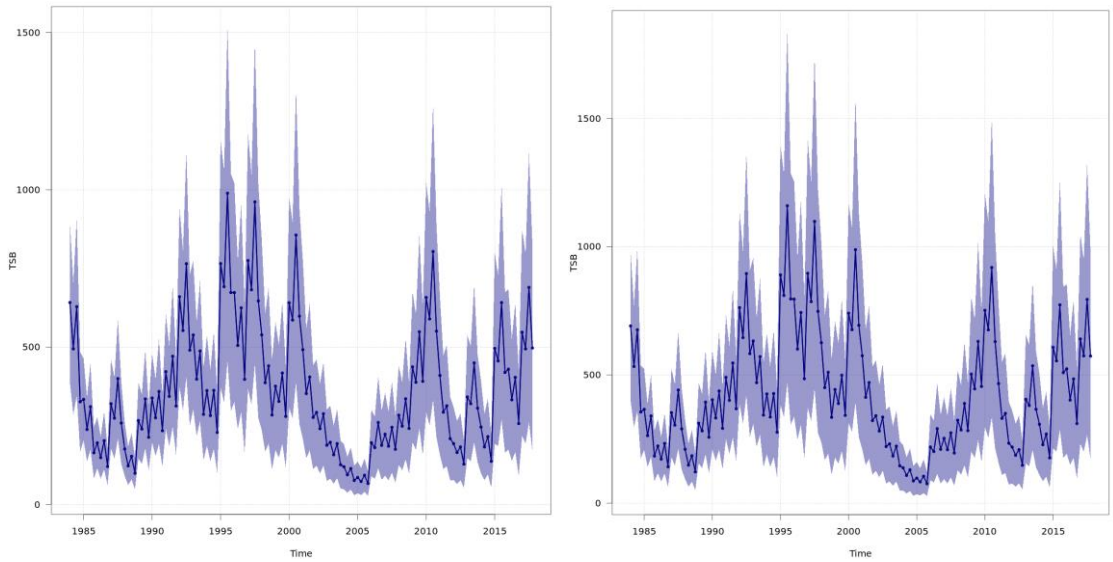

Figure A.2.2 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: TSB (t), quarterly. SESAM 2017 MSE Assessment.**
Old left; New right.

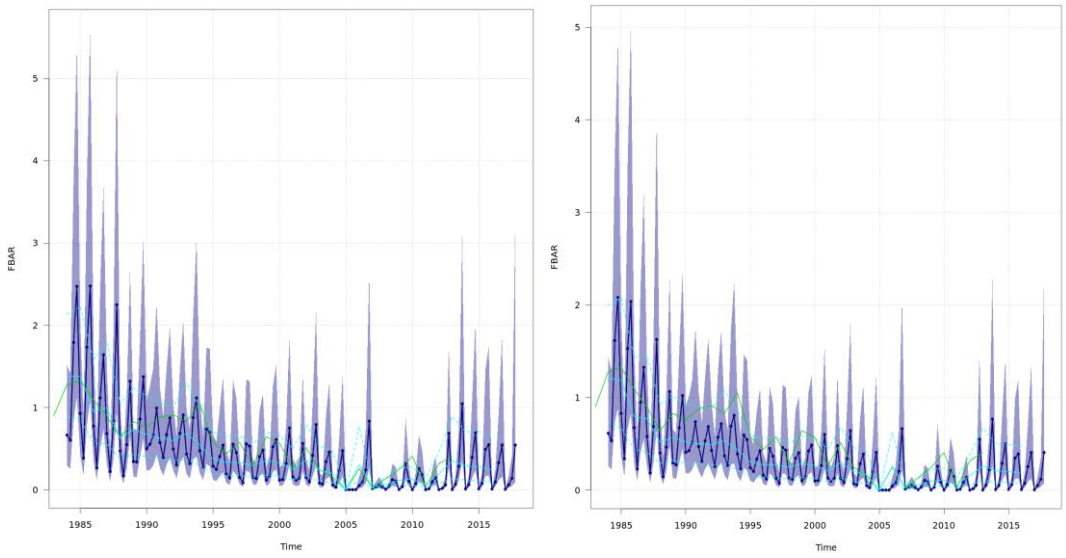

Figure A.2.3 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: $F_{1-2} = F_{bar}$, quarterly. SESAM 2017 MSE Assessment. Blue is quarterly values from SESAM, cyan is the yearly average from SESAM, green is yearly average from SXSA.**
Old left; New right.

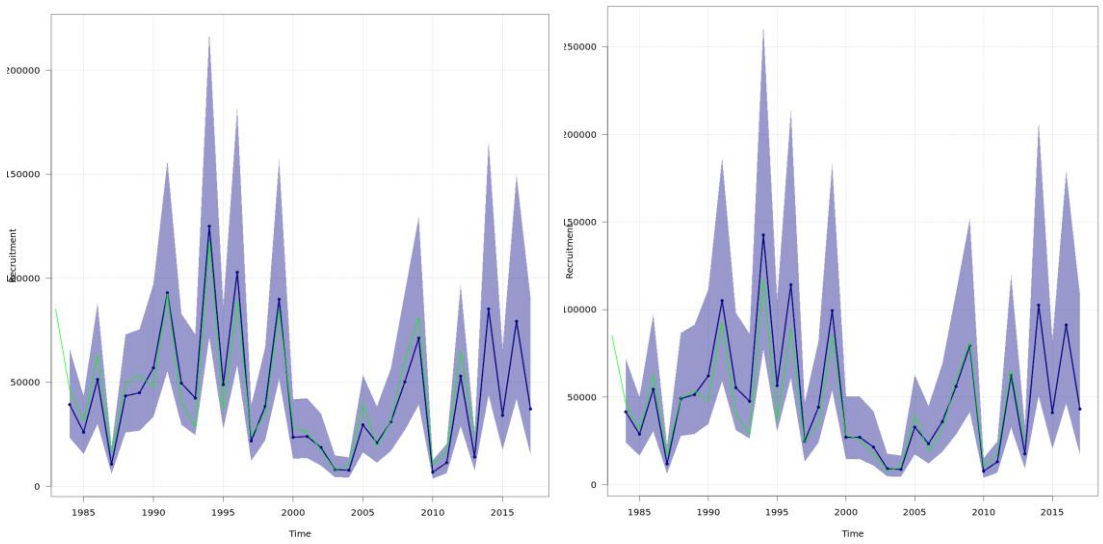

Figure A.2.4 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: Recruitment (millions), yearly. SESAM 2017 MSE Assessment. Blue is SESAM, green is SXSA.**
Old left; New right.

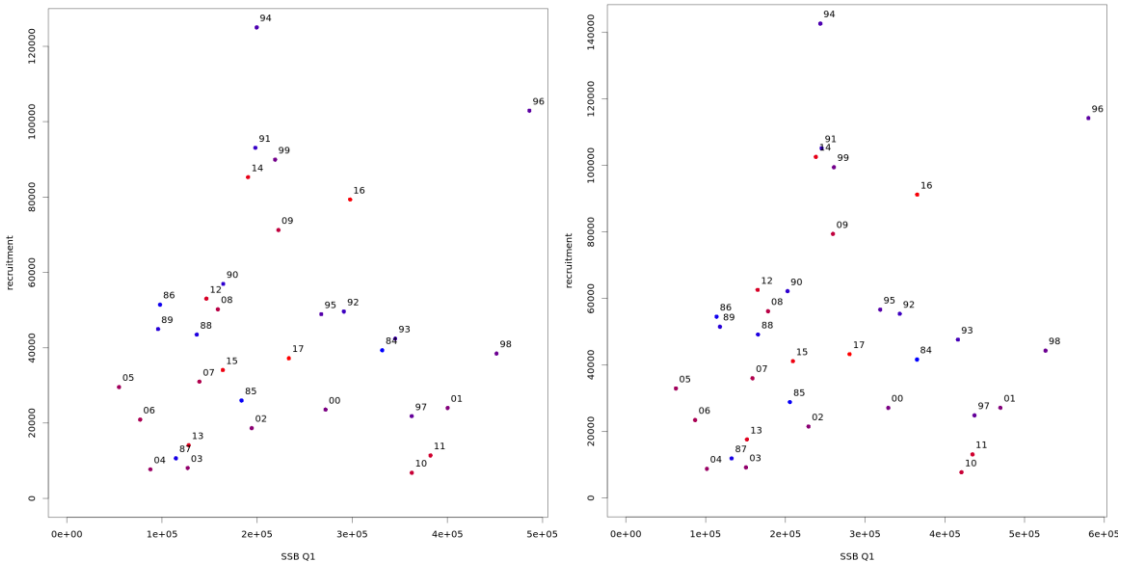

Figure A.2.5 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: Stock (SSB) – Recruitment Plot Quarter 1. SESAM 2017 MSE Assessment.**
Old left; New right.

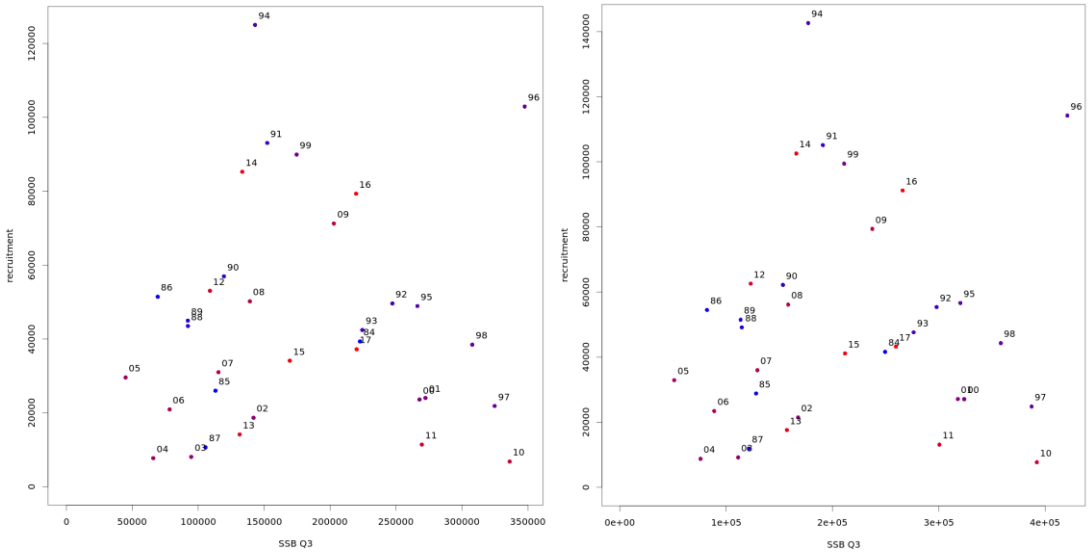

Figure A.2.6 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: Stock (SSB) – Recruitment Plot Quarter 3. SESAM 2017 MSE Assessment.**
Old left; New right.

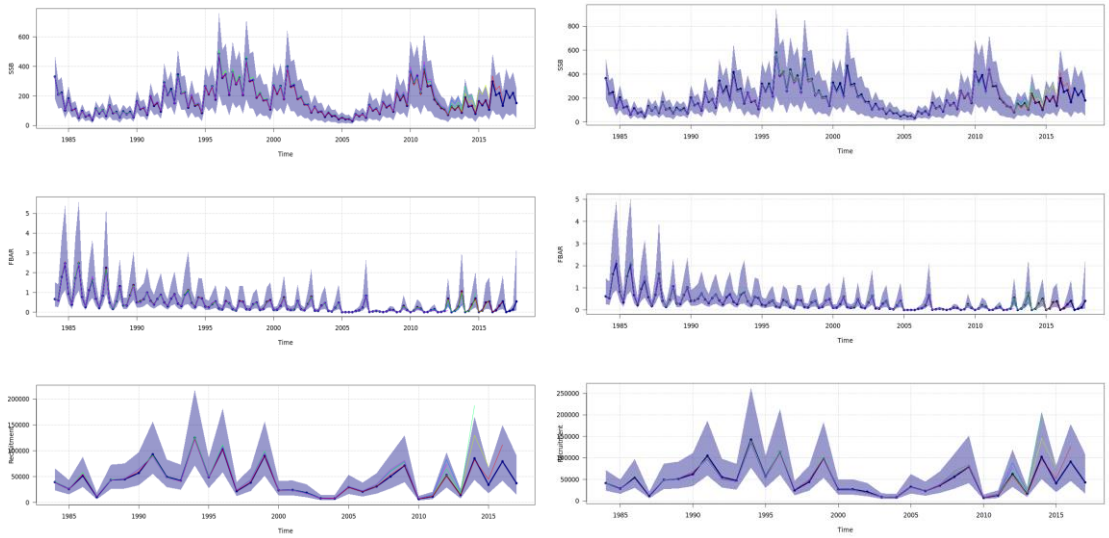

Figure A.2.7 **Norway pout in 4 and 3.aN (Skagerrak). Retrospective plots of baseline SESAM 2017 MSE Assessment, with terminal assessment year ranging from 2005–2017. Updated with Mohns Rho.**
Old left; New right.

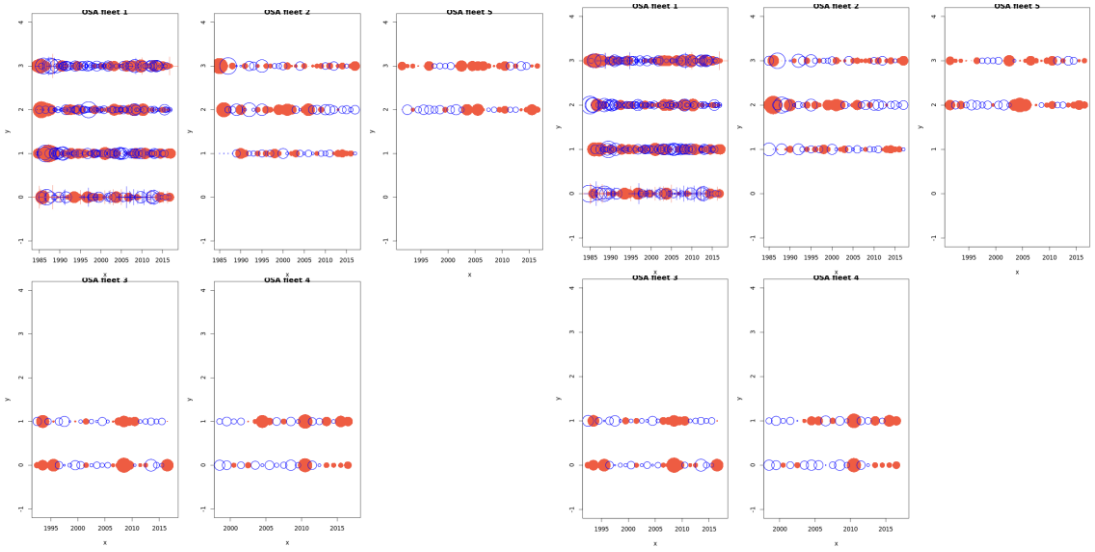

Figure A.2.8 **Norway Pout in 4 and 3.aN (Skagerrak). Assessment Diagnostics Plots by fleet and age: One step ahead residuals (see Berg and Nielsen, 2016). SESAM 2017 MSE Assessment. Old left; New right.**

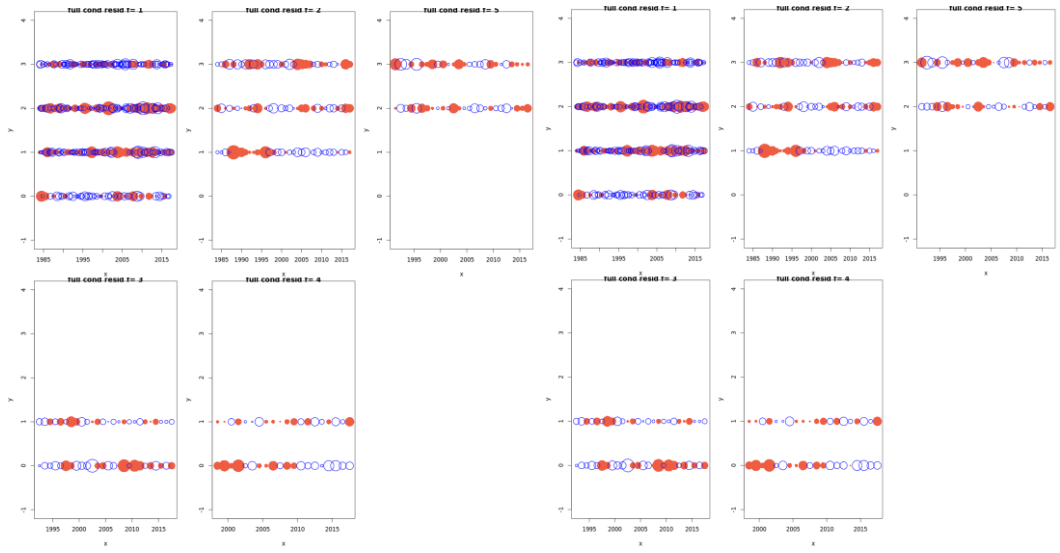

Figure A.2.9 **Norway Pout in 4 and 3.aN (Skagerrak). Assessment Diagnostics Plots: Full conditional residuals or auxiliary residuals by fleet and age (see Berg and Nielsen, 2016). SESAM 2017 MSE Assessment. Old left; New right.**

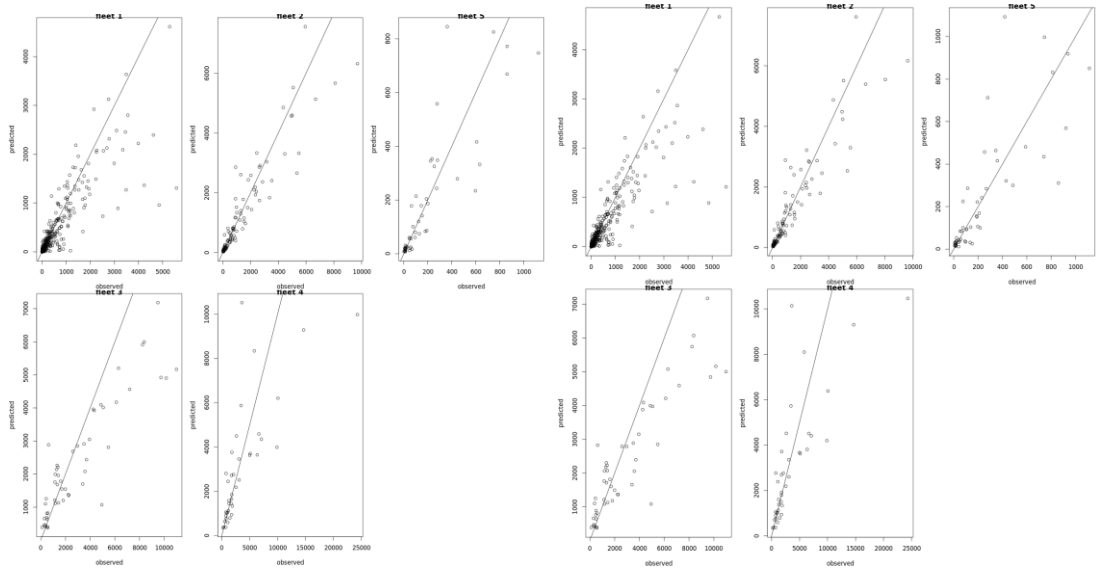

Figure A.2.10 **Norway Pout in 4 and 3.aN (Skagerrak). Assessment Diagnostics Predicted vs Observed catches. Plots by fleet. SESAM 2017 MSE Assessment.**
Old left; New right.

Appendix A.3.

Detailed assessment results from the 2016–2017 Benchmark Assessment with previous and revised IBTS data time series

Norway Pout 2016 Benchmark Assessment: Comparative assessment results between NorPoutBench2016 with previous IBTS data (left hand side plots) and NP_Sep17_fixC_Benchmark2016Data_NewIBTS with new IBTS data (right hand side plots). Summary of results – quarterly with uncertainties.

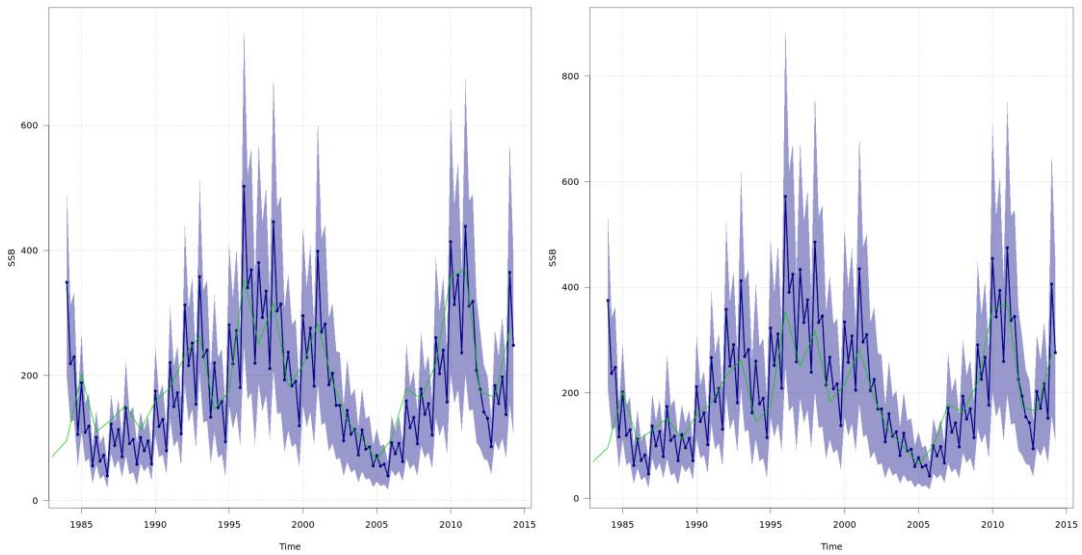

Figure A.3.1 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: SSB (t), quarterly. SESAM 2016-2017 Benchmark run. Quarterly estimated SSB and confidence interval from SESAM (blue) and SXSA (green, quarter 1 only – connecting lines are interpolations).**
Old left; New right

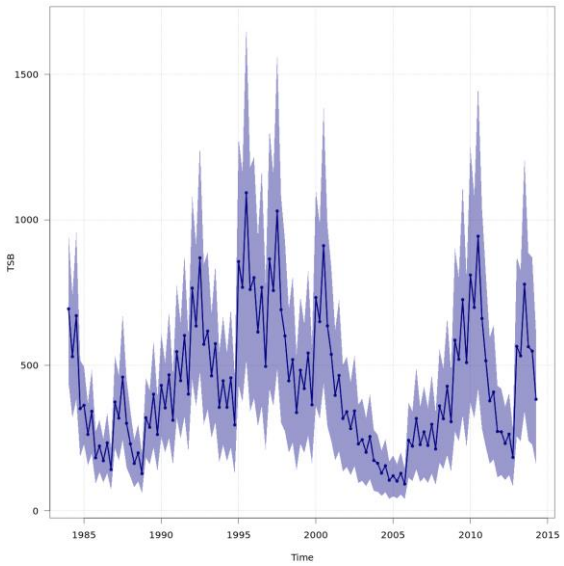

Figure A.3.2 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: TSB (t), quarterly. SESAM 2016–2017 Benchmark run. .**

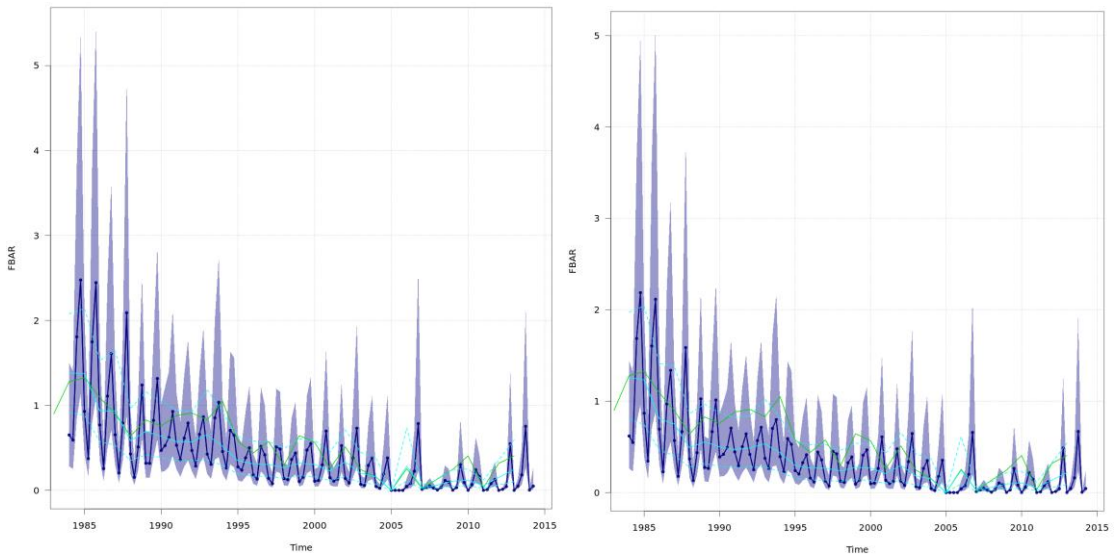

Figure A.3.3 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: $F_{1-2} = F_{bar}$, quarterly. SESAM 2016–2017 Benchmark run. Blue is quarterly values from SESAM, cyan is the yearly average from SESAM, green is yearly average from SXSA.**
Old left; New right.

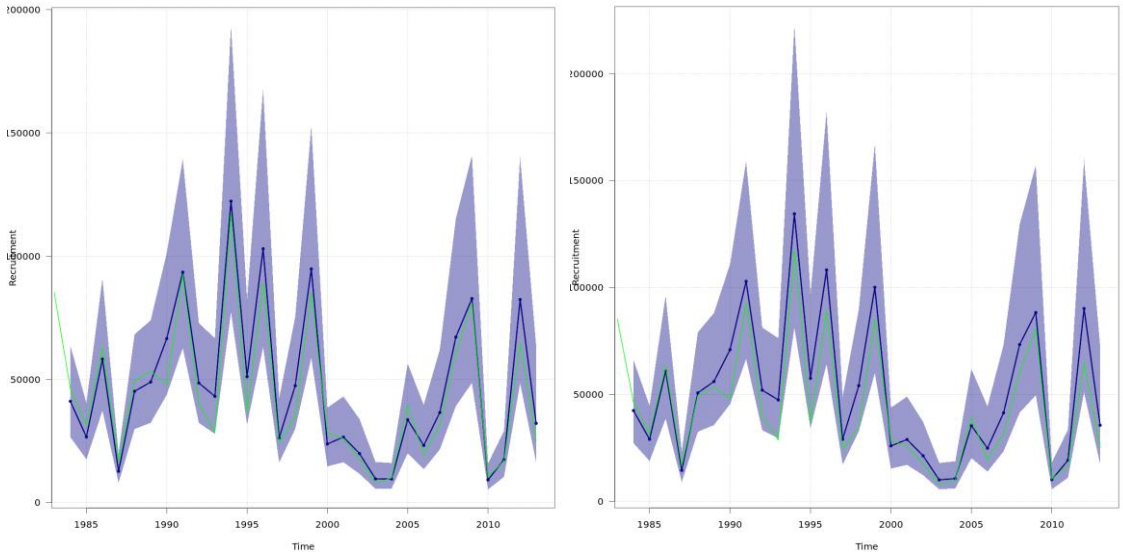

Figure A.3.4 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: Recruitment (millions), yearly. SESAM 2016–2017 Benchmark run. Blue is SESAM, green is SXSA. Old left; New right.**

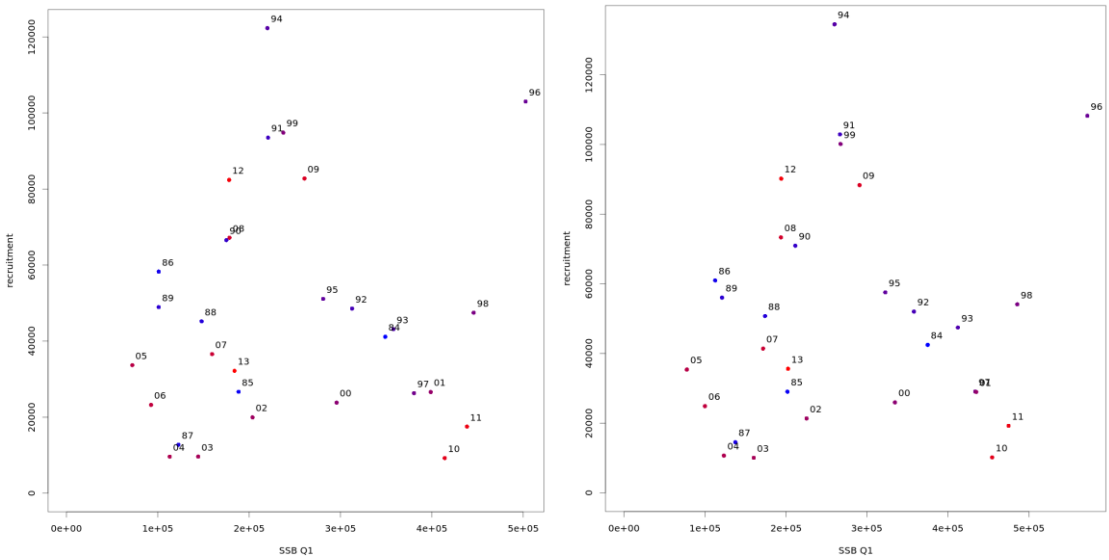

Figure A.3.5 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: Stock (SSB) – Recruitment Plot Quarter 1. SESAM 2016–2017 Benchmark run. Old left; New right.**

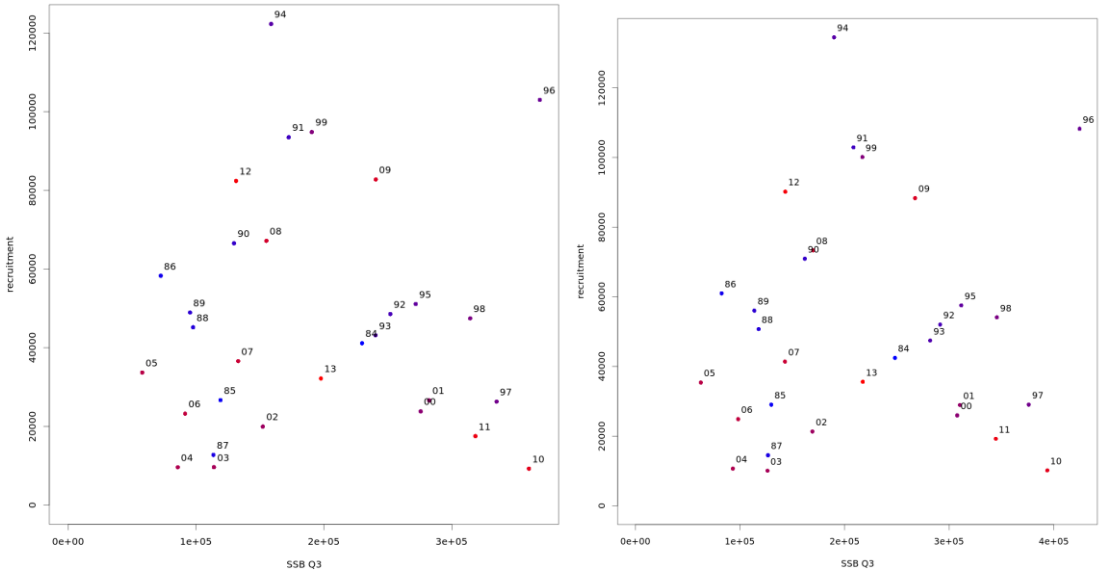

Figure A.3.6 **Norway Pout in 4 and 3.aN (Skagerrak). Stock Summary Plots: Stock (SSB) – Recruitment Plot Quarter 3. SESAM 2016–2017 Benchmark run. Old left; New right.**

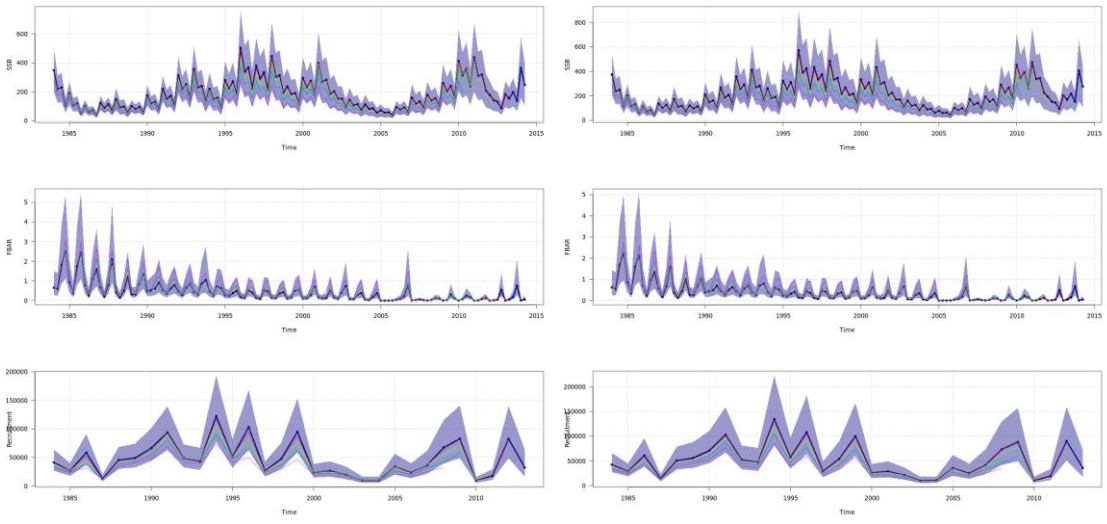

Figure A.3.7 **Norway pout in 4 and 3.aN (Skagerrak). Retrospective plots of baseline SESAM benchmark 2016–2017, with terminal assessment year ranging from 2005–2015. Updated with Mohns Rho. Old left; New right.**

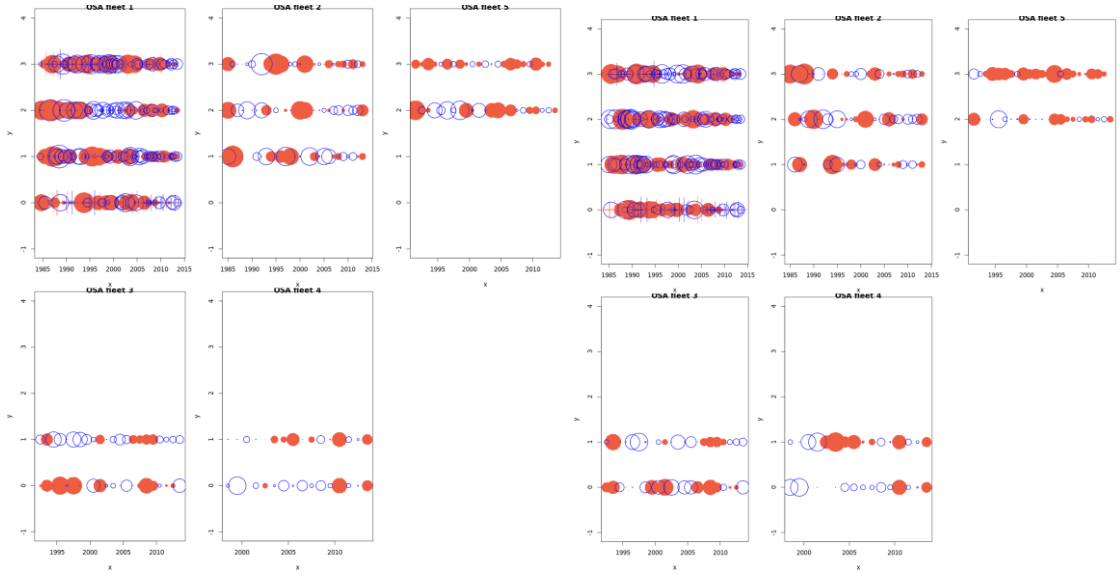

Figure A.3.8 **Norway Pout in 4 and 3.aN (Skagerrak). Assessment Diagnostics Plots by fleet and age: One step ahead residuals (see Berg and Nielsen, 2016). SESAM 2016–2017 Benchmark run. Old left; New right.**

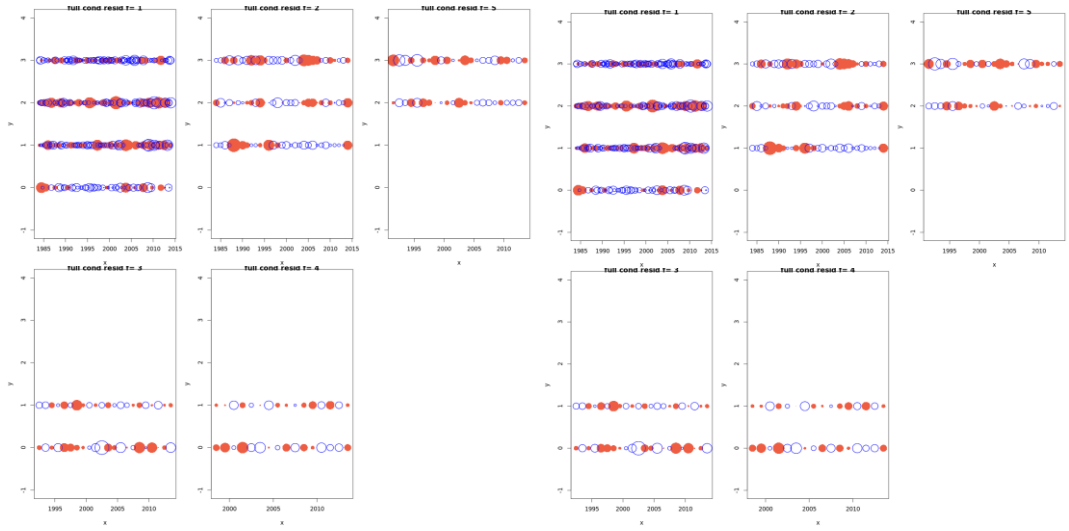

Figure A.3.9 **Norway Pout in 4 and 3.aN (Skagerrak). Assessment Diagnostics Plots: Full conditional residuals or auxiliary residuals by fleet and age (see Berg and Nielsen, 2016). SESAM 2016–2017 Benchmark run. Old left; New right.**

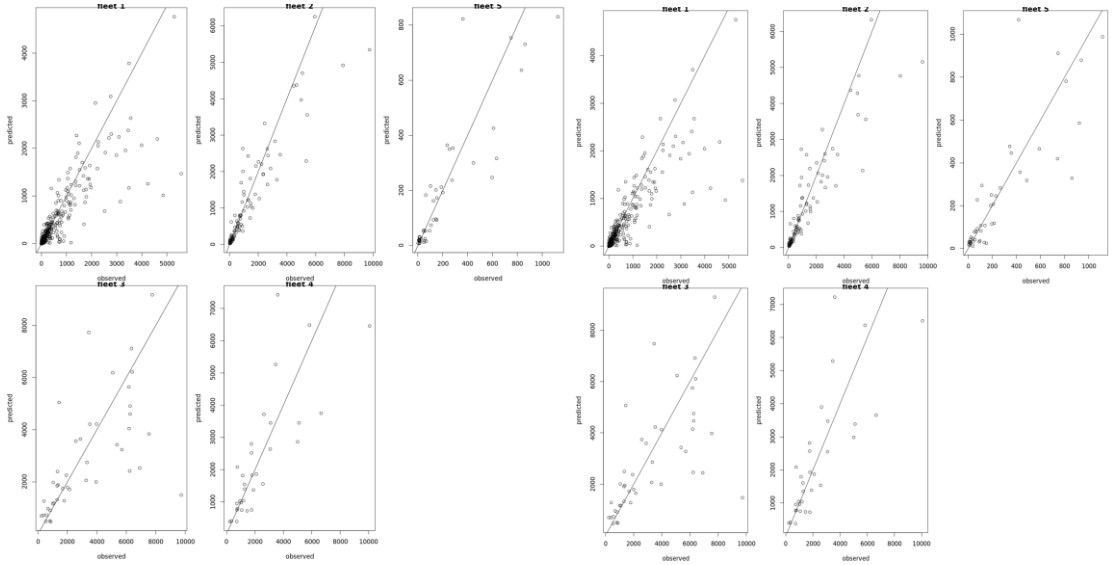

Figure A.3.10 **Norway Pout in 4 and 3.aN (Skagerrak). Assessment Diagnostics Predicted vs Observed catches. Plots by fleet. SESAM 2016–2017 Benchmark run.**
Old left; New right.

Appendix A.4

Extended MSE Results

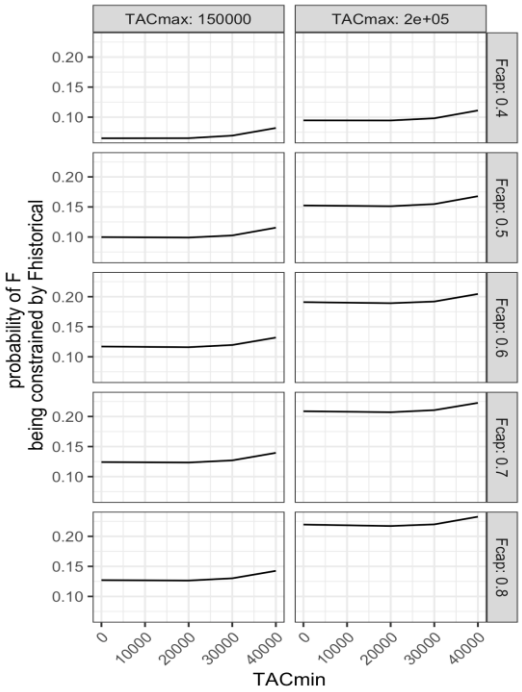

Figure A.4.1 **Probability of implementation error in main simulation results. This is the probability that $F_{historical}$ acts as an upper limit on implemented F_{bar}. This is "asF_{hist}" as defined in Table 5.**

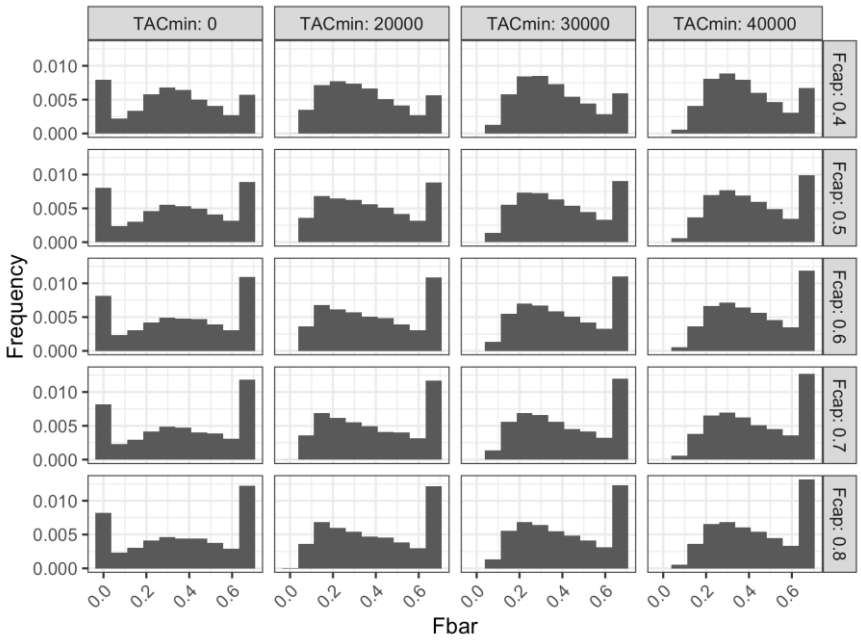

Figure A.4.2 **Frequency distribution across replicates and years 2023 to 2037 of F_{bar} depending on HCR with TAC_{max} = 200k tonnes.**

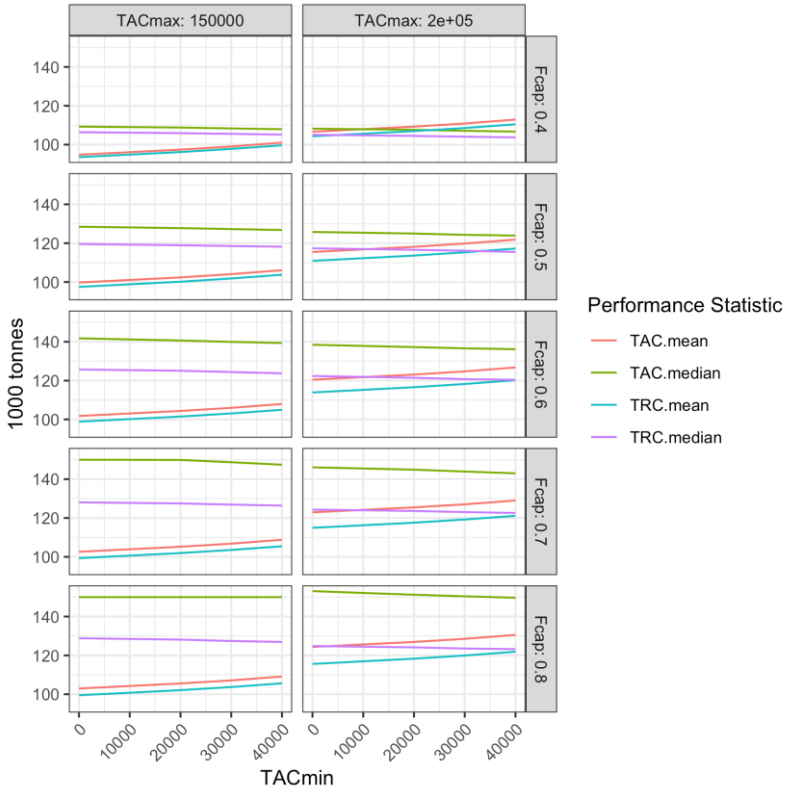

Figure A.4.3 **Performance statistics related to TAC as described in Table 5. TRC is total realized catch which may differ from TAC due to implementation error.**

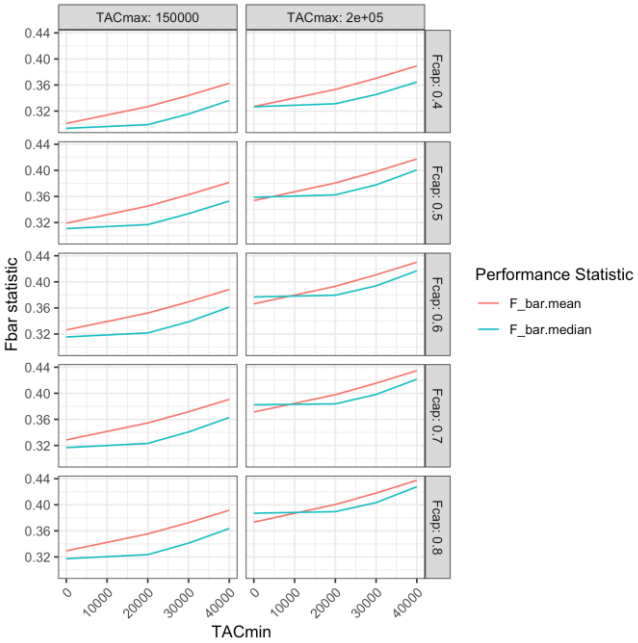

Figure A.4.4 **Performance statistics related to F as described in Table 5.**

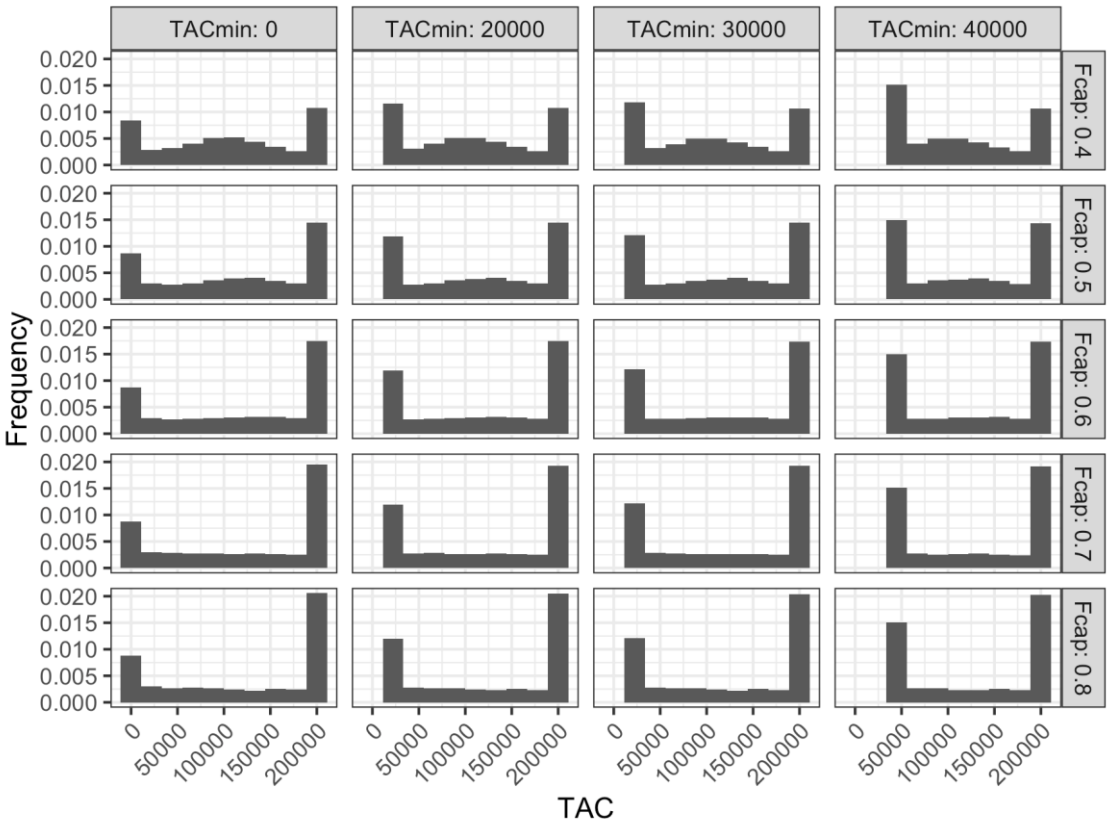

Figure A.4.5 **Frequency distribution across replicates and years 2023 to 2037 of TAC depending on HCR with TAC_{max} = 200k tonnes.**

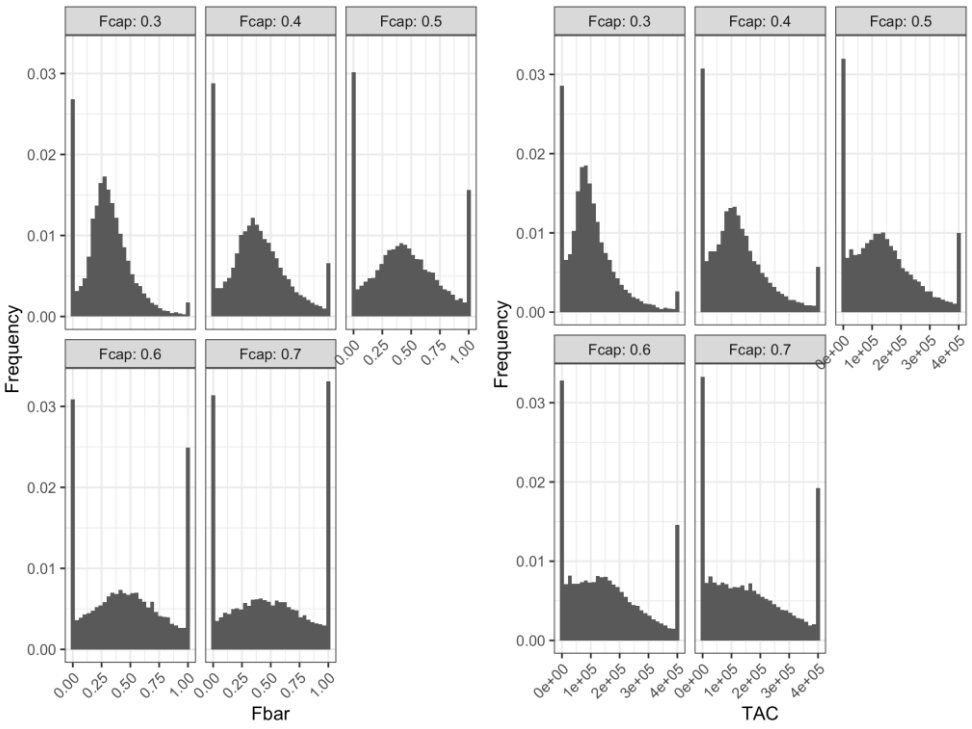

Figure A.4.6 **Sensitivity test results. Frequency distribution across replicates and years 2023 to 2037 of TAC (left) and F_{bar} (right) depending on Fcap with TAC_{min} = 0, TAC_{max} = 400k tonnes, and maximum realized F_{bar} = 1.**

Annex 1: Reviewers' comments

Review of Impact of revised IBTS survey indices on the Norway pout assessment out-put and sustainability reference points (Blim and Fcap), Parts I and II.

Reviewer 1

Responding to a request from ICES, I have briefly reviewed an analysis of the impacts of revised IBTS survey indices on the Norway Pout assessment output and reference points. Impacts on assessment output are relatively minor, though not negligible. The revisions to reference points were also minor and in the expected direction given the changes in assessment output. Rerunning the MSE evaluation indicated that the adopted management strategy for Norway pout would still be considered precautionary according to the ICES standard. My overall conclusion is that analysis was done appropriately and is suitable for providing management advice. I have several specific comments below that it would be good to address if possible in the final draft of the document.

Specific comments:

- It is stated that spawning stock biomass is estimated to be "slightly higher" in Figures 3-5, where a comparison is made between assessments using the old and revised IBTS survey indices. By my eye, it appears that SSB approximately doubled in Figures 3-5. Since recruitment estimates and total biomass estimates are about the same, I do not understand how SSB could change so much. I suspect that this is just a plotting error, since the percent differences in Blim estimates (SSB in 2005) in Table 4 are much smaller than in the figures.
- The value of Blim varies substantially between assessments in Table 4, with more recent assessments showing a lower Blim. The updated Blim is from a corrected 2017 MSE that uses updated data, but is still constrained to use only the data sets that went into the 2016 benchmark assessment. It seems very odd to me to update the Blim using an outdated assessment with revised survey indices, but to exclude consideration of more recent information. Is there a rationale for why this a good approach?
- It is stated that SSB in Q4 was used to calculate risk because Q4 is when the escapement strategy targets SSB. However my experience is that spawning stock biomass is always defined as the spawning stock at time of spawning, which appears to be Q1 for Norway pout. The MSE should evaluate risk using a SSB and a Blim in Q1. This distinction may not matter much, and the use of Q4 SSB to evaluate risk seems to be feature of both the old and updated analyses, so this could not be considered a concern that is unique to the updated analysis.

Reviewer 2

A long-term management strategy for providing TAC advice for Norway pout, based on an escapement strategy, was evaluated in September 2018. Several HCRs were simulation tested considering combinations of TAC_{min}, TAC_{max} and F_{cap} ($F_{bar(1-2)}$). The evaluations were conditioned by considering that the full TAC would not be taken if the required F to catch the TAC exceeded an upper F ($F_{historical}$ of 0.89, the 97.5 percentile of the stock assessment Fbar in the last 20 years). It was concluded that the escapement strategy for providing TAC advice for Norway pout was only precautionary by setting an F_{cap}=0.7 (no more than 5% probability of the spawning-stock biomass by 1 October in the forecast year falling below B_{lim}). At the time, B_{lim} (the lowest observed biomass in the beginning of Q4 in 2005 as estimated in the 2016 benchmark) was estimated to be 39 447 t and B_{pa} (=B_{lim} $e^{0.3 \times 1.645}$) to be 64 616 t.

The present review is based on the analysis and results presented in the following two working documents:

WD1 - Impact of revised IBTS survey indices on the Norway pout assessment output and sustainability reference points (Blim and Fcap), J. Rasmus Nielsen and Mollie Brooks, DTU Aqua

WD2 - Impact of revised IBTS survey indices on the Norway pout assessment output and sustainability reference points (Blim and Fcap) – Part II, Mollie Brooks and J. Rasmus Nielsen, DTU Aqua

My comments addresses three aspects, 1) the impact on the stock assessment of Norway pout due to the revision of the time series of IBTS Q1 and Q4 survey indices, 2) the re-evaluation of the Norway pout B_{lim} and, 3) the re-evaluation of HCR with F_{cap}=0.7.

1) Impact on the stock assessment of Norway pout due to the revision of the time series of IBTS Q1 and Q4 survey indices

WD1 does not provide information on the reason behind the revised IBTS Q1 and Q4 survey indices for Norway pout. Therefore my comments rely on the assumption that an internal WGNSSK review was carried out and that the group concluded that the revision of the IBTS Q1 and Q4 survey time series resulted in an improvement in the quality and reliability of the abundance-at-age time series used in the stock assessment of Norway pout.

There are substantial differences between the old and the new indices, particularly after 1991 in ages in 3-4 in Q1 and in ages 2-3 in Q3 (Tables 1 & 2 and Figures 1 & 2 in WD1) with the majority of the new values resulting in higher abundance indices. These age classes contribute 100% to SSB, hence influencing SSB.

The impact of the new IBTS survey indices on the Norway pout assessment was conducted by comparing the estimates of the stock key parameters (R, SSB, Fbar) with the stock assessment estimates using the previous (old) IBTS survey indices. For the comparisons conducted I focused on the results from the assessments using the new data: the 2019 stock assessment (in March 2020), the 2017 MSE assessment (SESAM run with bug code corrected) and the 2017 MSE with the 2016 benchmark data (used as basis for the MSE simulation testing of the management strategy for Norway pout conducted in Sep 2018).

The revised IBTS survey indices (New) resulted in a scaled revision of SSB and F_{bar} over the time series: SSB upward and F_{bar} downward. Recruitment was slightly revised upward. Tables with the old and the new stock assessment estimates were not provided but a visual inspection of Figures 4 (2019 assessment) and Figure 5 (2017 assessment) indicate that the magnitude of the upward revision of SSB was higher in the periods 1996-2001 and 2009-2011. The CIs of the new and the old estimates overlap, hence I agree that the performance of the assessments and stock trajectories are approximately identical.

These results prompted the revision of B_{lim} and B_{pa} and to re-run the MSE to evaluate whether a management strategy for Norway pout with an F_{cap}=0.7 is still precautionary.

2) Revised BRPs B_{lim} and B_{pa}

The new S-R scatter plots do not change the rationale to derive B_{lim} as the lowest observed biomass in the beginning of Q4 in 2005. B_{lim} and its 95% CI with the revised (new) IBTS abundance time series are:

2019 (March 2020) stock assessment – B_{lim}=**32 456 t** (6 970 t, 57 941 t)

2017 MSE assessment – B_{lim}= **34 999 t** (11 013 t, 58 985 t)

2017 MSE with 2016 benchmark data– B_{lim}= **42 573 t** (17 046 t, 68 100 t)

In comparison to the value adopted in the 2016 benchmark assessment (17 736 t, **39 447** t, 61 158 t), the point estimates of B_{lim} are revised downward with the '2019 March stock assessment' (18%) and with the '2017 MSE assessment' (11%) and upward when running the benchmark assessment with the new indices, i.e, the '2017 MSE with the 2016 benchmark data' (8%). B_{lim} estimated from the '2017 MSE with 2016 benchmark data' is slightly higher than the value previously adopted and the CIs of the old and the new B_{lim} estimates overlap, indicating that there is no statistically significant difference between estimates. Therefore, I agree with the group to adopt B_{lim} = 42 573 t, hence B_{pa}= 69 736 t.

3) Re-evaluation of F_{cap}=0.7

MSE simulation testing was performed for a set of HCRs: combinations of TAC_{max} (150 000 t & 200 000 t), TAC_{min} (0 to 40 000 t, at interval 10 000 t) and F_{cap} (0.4 to 0.8, at interval 0.1). The MSE OM was conditioned using the estimates from the 2017 assessment with the new IBTS time series and followed the same assumptions (e.g. R governed by a hockey-stick model with an inflection point at 62 683 t - SSB estimated in Q1 in 2005;) and approach (e.g. using a stock assessment emulator to mimic the SESAM assessment in the MP component) as in the 2018 benchmark MSE. Implementation error was introduced by assuming that F_{bar} would not exceed 0.67, the $F_{historical}$ corresponding to the maximum 97.5% percentile of F_{bar} estimated in the last 20 years of the 2017 assessment (run with new IBTS indices). The performance of HCR assuming implementation error was tested for the options TAC_{min}=0 t, TAC_{max}=200 000 t and F_{cap} in the range 0.3-0.7. Sensitivity analysis was carried out to evaluate the performance of HCRs with F_{cap} in the range 0.5-0.8 but with no limits in TAC_{max} and allowing F_{bar} to increase up to a maximum of 0.67, 1.0 and 2.0. Simulations were based on 1000 replicates, projected from 2018 to 2037.

The P(SSB < B_{lim}=42 573 t in Q4) was evaluated against risks type 1 and type 3 in Q1 and Q4, computed for the short-term (2018-2022) and the long-term (2023-2037). F_{cap}=0.7 is considered precautionary in HCRs with TAC_{min} of 0 t, 10 000 t, 20 000 t and 30 000 t and TAC_{max} of 150 000 t and 200 000 t: the P(SSB < B_{lim}) is below 5% for risk type 1 and type 3 in both quarters and in the short- and long-term. The results from the sensitivity analysis show that F_{cap}=0.7 is still precautionary in a scenario allowing F_{bar} to increase to F_{max}=1.0 though with a high probability of having closures in the fishery.

Technical comments and recommendations, provided during the review of the 2018 workshop for the management strategy for Norway pout still apply. However, despite a 'short-cut' approach used in the current simulation testing, considerable exploratory analysis were conducted. Risks type 1 and type 3 computed for the HCR with F_{cap}=0.7, TAC_{min}=0 t, maximum TAC capped at 200 000 t and maximum F capped at 0.67 are well below 5% and, still below 5% if F_{bar} is allowed to increase to 1.0 giving confidence to the conclusion that F_{cap}=0.7 is precautionary.

Please note the following:

- Figure 7 of WD1 only shows risk values for F_{cap} range 0.4-0.8;
- Figures A.2.5 & A.2.6 (Annex WD1) show the scatter plot of SSB-R (SESAM 2017 assessment estimates; old *vs* new indices). It looks like the blue dots represent the SSB-R estimates in the earlier years and the red dots the estimated for the most recent period. Please clarify.
- For an F_{cap}=0.7 and assuming F_{max}=2.0, both risk3.short.Q4 & risk3.long.Q4 are above 5% (WD2 - Fig 4.1; see below) but in the 'Conclusions' section it is stated that 'even with the unrealistically high maximum implementable F, then the only risk that goes above 0.05 (when rounded to the nearest 0.01 units) is risk3.long.Q4 for a Fcap=0,7.'