ICES WKMACQI REPORT 2018

ECOSYSTEM OBSERVATION STEERING GROUP

ICES CM 2018/EOSG 34

REF. ACOM & SCICOM

Report of the Workshop on Mackerel biological parameter Quality Indicators (WKMACQI)

15-17 May 2018

IJmuiden, The Netherlands

ICES
CIEM
International Council for the Exploration of the Sea
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

Recommended format for purposes of citation:

ICES. 2018. Report of the Workshop on Mackerel biological parameter Quality Indicators (WKMACQI), 15–17 May 2018 , IJmuiden, The Netherlands. ICES CM 2018/EOSG 34. 42 pp.

Contents

Executive summary

The Workshop on Mackerel biological Quality Indicators (WKMACQI) met 15–17 May 2018 in IJmuiden, The Netherlands. WKMACQI is initiated by the Working Group on Biological Parameters (WGBIOP) and aimed to carry out sensitivity analyses of the mackerel assessment for uncertainty in biological parameters.

Many biological parameters are collected for assessment purposes, but quality indicators are rarely available for these parameters. In those few cases, when they are available, the quality indicators are not incorporated in the assessment process. In the past three years, WGBIOP has developed qualitative and quantitative quality indicators for biological parameters. Ambitiously, WGBIOP wanted to incorporate quality indicators in the assessment process, but this goal has not been reached up till now. This was due to the fact that it was not possible to get stock assessors involved in the WGBIOP meetings. WGBIOP did receive positive reactions on the work that was carried out on the quality indicators and this workshop is the result.

For mackerel, age and maturity data were available to develop a quantitative quality indicators. First the data available to estimate the maturity ogive was analysed to check for outliers and if there is evidence of sexual dimorphism in mackerel. Some 'older' mackerel (5+), mostly females, were noted as immature. This is probably due to the fact that mackerel has a long spawning season. Large females start spawning early and are already in spent or resting stage while younger fish are still spawning. These spent or resting females are then macroscopically easily confused with immatures. Despite these outliers the resulting maturity ogive seems reasonable and there is no evidence of sexual dimorphism.

Age and maturity calibration exercises have been carried out in the recent past. Data of these calibrations provide an uncertainty measure of ageing and maturity staging. Age (AEM) and Maturity Staging (MSEM) Error Matrices were developed. The error matrix gives the probabilities that a sampled fish of true age/maturity class *a* is assigned to one of the observed age/maturity classes. For age the 'true age' is based on the modal age. Maturity staging can be more easily validated with the use of histology, thus the maturity stages are checked against the 'true maturity'.

WKMACQI made the assumption that the data available on the mackerel stock is not affected by any error on age or maturity stage determination, and the WGWIDE 2017 data and assessment are the reference. The error matrices are used to "pollute" this input data and the assessment is run on the polluted data to determine the sensitivity. Each of the different parts of the assessment model, were age and maturity are used, were investigated separately and after that an assessment with all data combined was carried out.

The analyses show that errors in the determination of biological parameters affected the mackerel assessment at different levels and can have a substantial effect on the output of the assessment. For instance, when ageing errors are affecting all data sources, a difference of +14% in the SSB and –14% in F_{bar} is observed. This is a substantial difference, and the SSB and F_{bar} trends of the model based on data affected by ageing errors are close to the limit of the confidence bounds of the WGWIDE 2017 assessment. Also the weighting of the different input data sources is affected, which seems to have different consequences depending on the assessment model used. The

sensitivity of assessment methods to these errors should be investigated in a more systematic way to understand the model and species-specific consequences of these errors.

WKMACQI was an excellent opportunity to work on biological data with people from both the assessment and biology side. This increased the understanding of the uncertainties in biological data, how biological data are used in assessments, and what data are needed to evaluate the effect of uncertainties on the outcome of the assessment. Such a close collaboration cannot be achieved by assessors participating in WGBIOP or survey people participating in the assessment group.

1 Introduction

The Workshop on Mackerel biological Quality Indicators (WKMACQI) met 15–17 May 2018 in IJmuiden, The Netherlands. Formally 7 participants from 4 countries participated in the meeting (Annex 1), but 3 could only participate via correspondence and by reviewing the report.

Terms of reference for the meeting are:

a) Review and consider quality indicators for and issues with biological parameters of western, southern and North Sea mackerel;

b) Prepare and update the Age Error Matrix and Maturity Staging Error Matrix;

c) Carry out sensitivity analyses of the mackerel assessment with regards to the quality indicators of mackerel biological parameters.

Many biological parameters of fish are collected for assessment purposes, but quality indicators are rarely available for these parameters. And if quality indicators are available, these are not included in the assessment model or considered in the assessment process. WGBIOP developed guidelines for qualitative and quantitative quality indicators for various parameters (ICES 2017a). Furthermore WGBIOP has been evaluating issues with biological parameters for stocks which have a benchmark coming up (ICES 2016, 2017a). But the goal to incorporate quality indicators in the actual assessment has not been reached within the first term of WGBIOP. This was due to the fact that WGBIOP has not been able to get stock assessors involved in the WGBIOP meetings. Nevertheless, contact has been established with the stock coordinators and issues and quality indicators on biological parameters have been put forward to them. WGBIOP received positive reactions from the stock coordinators which facilitated a qualitative consideration of the issues in the assessment process (ICES 2017a). WGBIOP still underlines the necessity to improve the assessment process by including quality indicators. Thus WGBIOP initiated WKMACQI with mackerel as a case study.

2 Adoption of the agenda

The agenda addressed all ToRs and can be found in Annex 2. The meeting started with an introduction to the workshop and presentations on age error estimations and sensitivity analyses of the mackerel assessment to ensure all participants received the necessary back ground information for the workshop. This was followed by a plenary discussion to decide which biological parameters were important for the mackerel and for which we could get quality indicators. Based on this, tasks were assigned to all participants. The workshop continued with practical work on the various tasks and regular plenary discussions of preliminary results.

3 Mackerel

Northeast Atlantic (NEA) mackerel is a wide-ranging and important commercial fish species. Mackerel spawn from the end of January until July in the Northeast Atlantic. Spawning areas range from west of Portugal to north of Scotland. The summer feeding grounds are found in the Nordic seas and adjacent areas, between Scotland, around Iceland and Norway. Mackerel thus make extensive migrations between the spawning and summer feeding grounds.

In the recent past the Northeast Atlantic mackerel stock has increased in size (ICES 2017b) and due to this both the spawning and summer feeding ground have increased in size. This increase in the last decade in the size of the mackerel stock has increased the number of countries fishing mackerel (Figure 3.1). Especially the northern European (UK, Norway, Iceland and the Faroe islands) countries have high reported mackerel catches.

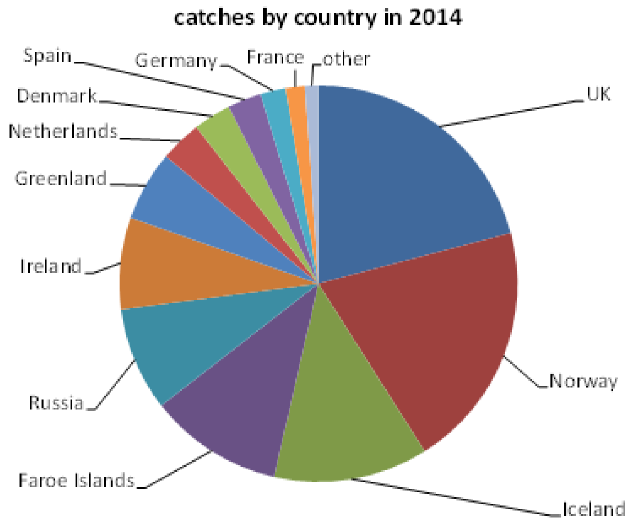

Figure 3.1. Mackerel catches by country for 2014 (as an example).

In Northeast Atlantic mackerel three different components are recognised; western, southern and North Sea. However, Northeast Atlantic mackerel is assessed and managed as one stock. For the assessment data from commercial catches are used as well as fisheries independent data from the Mackerel and Horse mackerel egg survey (MEGS), North Sea Mackerel egg survey (NSMEGS), demersal trawl surveys (IBTS Q1 & Q4), ecosystem surveys in the Nordic Seas (IESSNS) and tagging data (ICES 2017b). Except for the egg surveys, the data from catches and surveys are from Northeast Atlantic mackerel, without separating for the different stocks. The egg surveys provide data on the different components separately (ICES 2017c, 2018).

4 Mackerel quality indicators for biological parameters (ToR a)

Biological data used in the mackerel assessment are length, weight, age, sex, maturity, natural mortality and stock structure. Within the egg surveys also egg staging and fecundity are biological parameters used to estimate the survey index. These biological parameters are associated with an uncertainty in the estimation. This uncertainty can come from various sources and may have a low or high impact on the stock assessment.

Unfortunately, for most biological parameters, insufficient data are available to investigate uncertainty and to produce quality indicators. Exceptions are age and maturity and these parameters were examined during WKMACQI.

Uncertainty in age and in maturity data can be quantified by an error matrix. This quantitative quality indicator is calculated for age in section 4.1 and for maturity in section 4.2. The maturity data were further explored to assess the main spawning season, scrutinise the data and examine sexual dimorphism in age at length.

IT was decided to focus on biological parameters which are used directly in the assessment of mackerel, due to time constraints. Thus workshop did not to carry out a further extensive review of quality indicators and issues with biological parameters such as egg identification or staging. These biological parameters are used in the estimation of the index from the mackerel egg survey, and not in the assessment directly.

4.1 Mackerel maturity data

The maturity ogive used in the mackerel assessment is updated yearly. At WKMACQI, the available dataseries used to determine the proportion of mature fish has been analysed.

The mackerel maturity-at-age dataseries is available from 1960 until 2016. The data format contains the following variables: year, month, ICES area, country (providing the data), age, length, weight, sex and (macroscopic) maturity stage. The fish were classified in four maturity stages: immature (i), mature (m), running (r) and spent (s). According to these maturity stages the fish at immature stage were assigned as immature and the ones classified in the other stages were considered to be mature fish. The maturity staging of individuals was based on visual inspection of the gonads (i.e. macroscopic determination), thus these maturity stages have not been validated (by histology).

The data analysis on the mackerel maturity data has the following purposes:

- To assess the main spawning period (section 4.1.1);
- To scrutinise the dataset (section 4.1.2);
- To examine the existence of sexual dimorphism in age at length (section 4.1.3).

4.1.1 Main spawning period

In the mackerel stock assessment, the period from February to July is considered as the main spawning period. The data from this period are used for the construction of the maturity ogive. The presence of individuals at the running stage is higher between February and July (Figure 4.1.1.1), which is in accordance with the spawning period used to estimate the maturity ogive.

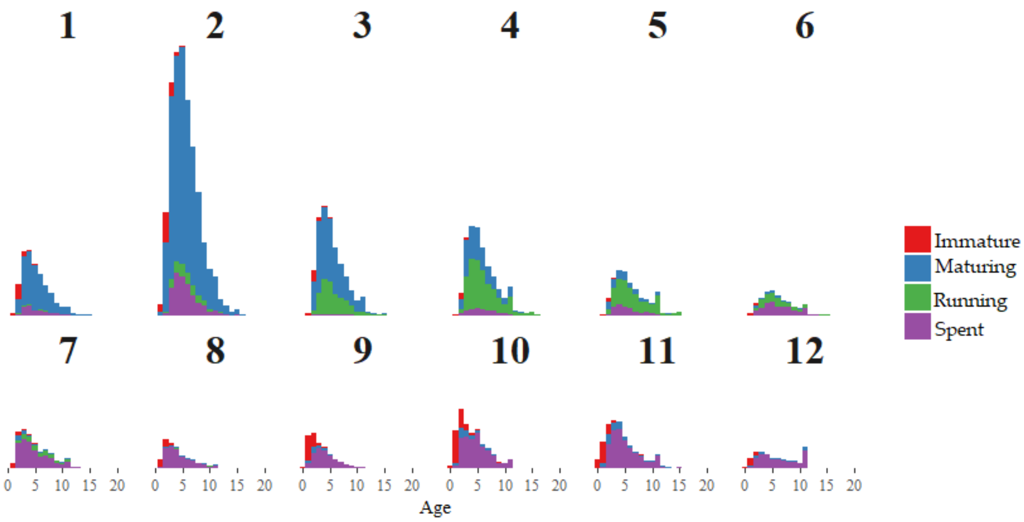

Figure 4.1.1.1. Distribution of the individuals by age group, maturity stage and month (from month 1 – January until month 12 – December).

4.1.2 Proportion of mature fish at age

The dataset available for determining the proportion of mature fish at age is explored in this subsection.

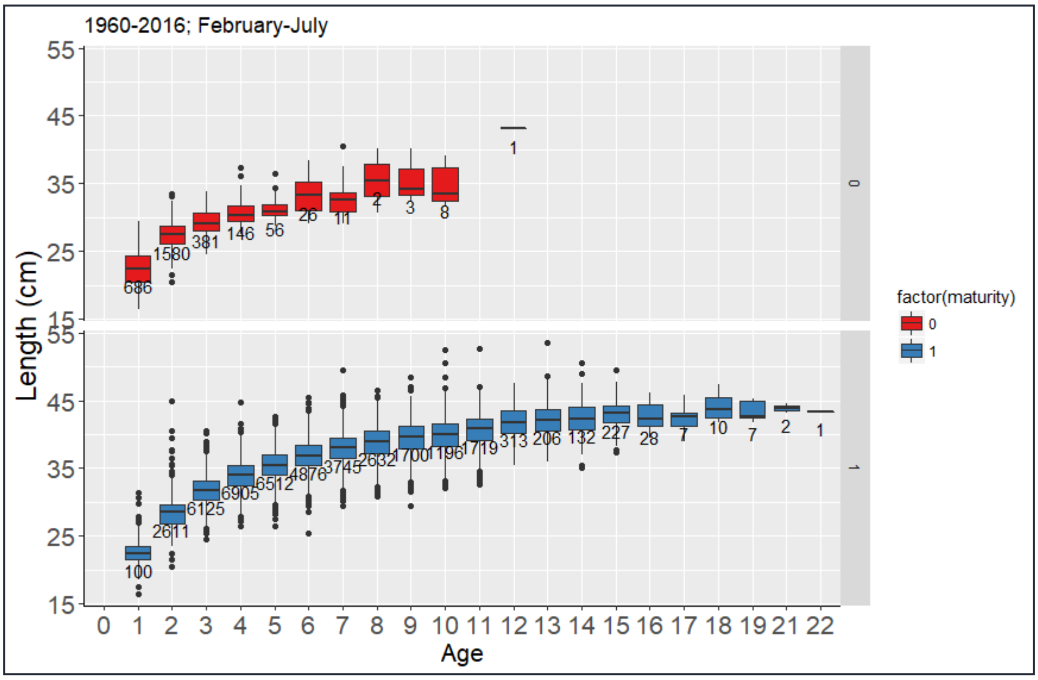

Figure 4.1.2.1. Length distribution of individuals by age group and by maturity (0 – immature; 1 – mature) from 1960 to 2016 and between February and July. The numbers below each boxplot indicate the number of individuals.

For the period 1960 until 2016, this dataset shows immature mackerel distributed in the age groups from 0 to 12 (Figure 4.1.2.1). This indicates uncertainties in the maturity staging classifications (immature stages *vs.* mature stages), because the presence of immatures was only expected until age 5. Upon closer examination, the age 6 and older mackerel which were staged immature, all are part of the historic Dutch data. At that time the Netherlands used the RIVO 8 maturity scale, in which both stage 1 and 2 were immature fish, but stage 2 already showed development of the gonads. All the older immature fish were in stage 2 in the RIVO 8 scale. There may have been an interpretation difference in the macroscopic maturity staging with just developing (true stage 2) and resting (different from stage 2). However, this can't be checked since the gonads were not preserved or photographed.

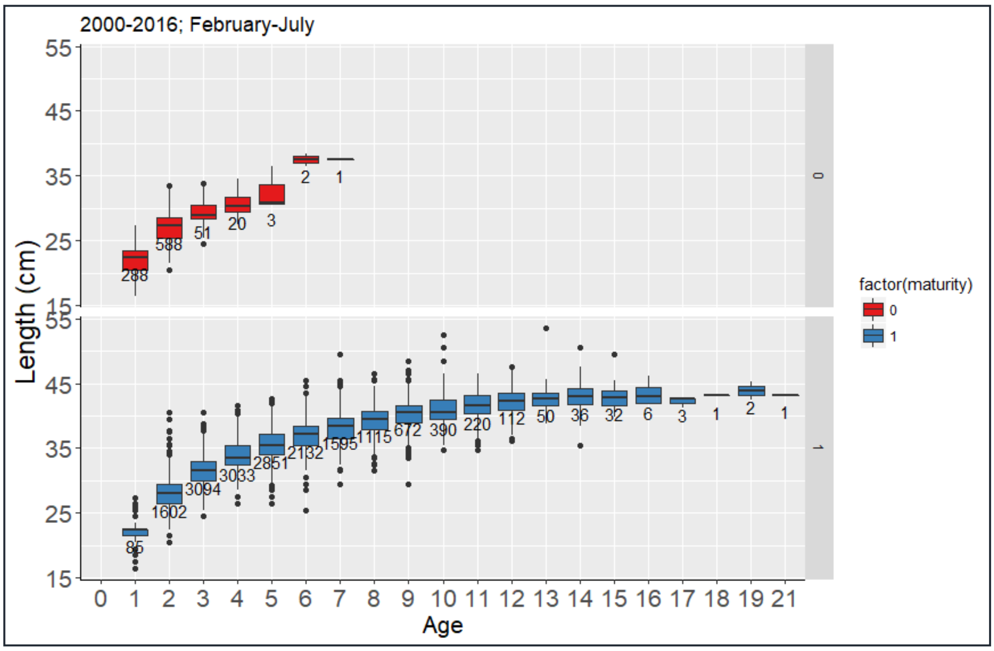

Figure 4.1.2.2. Length distribution of individuals by age group and by maturity (0 – immature; 1 – mature) from 2000 to 2016 and between February and July. The numbers below each boxplot indicate the number of individuals.

From 2000 to 2016, the number of immature mackerel at age 4+ is very small compared to the number of matures in the same age groups (Figure 4.1.2.2). Consequently, the proportion of matures is ~1 from age 4 onwards. Although, this doesn't affect the maturity ogive it indicates a possible misclassification between immature and mature stages. This misclassification is probably due to the similarities in the macroscopic appearance of gonads in the immature and spent stages. This type of misclassification, spent gonads classified as immature, increases the number of immatures and introduces errors in the maturity ogive.

During peak spawning numbers of spent mackerel are very low (unpublished data from ICES WGMEGS). One solution could be readjusting the period considered each year, to avoid this kind of errors on data. However, the solution which guarantees the most accurate maturity stage classification is the use of histology. To improve the accuracy of the data used for establishing the maturity ogive, a subsample of macroscopically classified immature gonads should be collected, by age group, to be processed histologically. It will then be possible to validate the macroscopic staging of older immature fish.

4.1.3 Sexual dimorphism in length-at-age

In species with sexual dimorphism in length-at-age, the number of individuals by sex used to construct the maturity ogives must be the same, to avoid biased data due to sampling (Gonçalves *et al.* 2017). The possible existence of different growth patterns and maturation between mackerel females and males was investigated (Figure 4.1.3.1 & 4.1.3.2). The age-length distribution for immature mackerel (Figure 4.1.3.1) and also mature mackerel (Figure 4.1.3.2) did not show a significant difference in growth between the sexes.

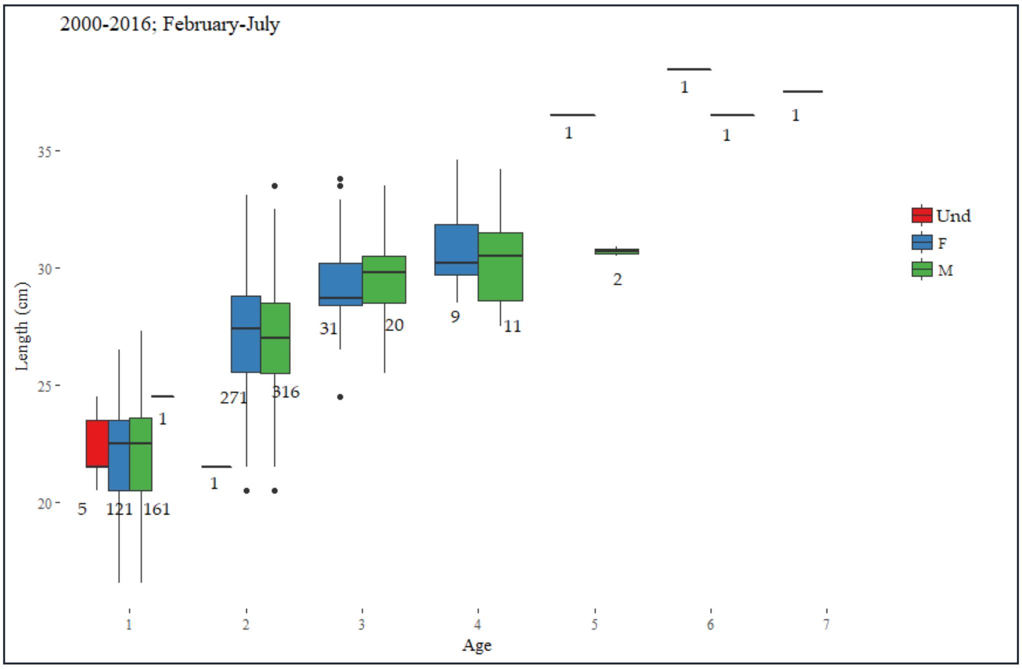

Figure 4.1.3.1. Length distribution by age group and sex for immatures (Und- undetermined sex; F – females; M – males), from 2000 to 2016 and between February and July. The numbers below each boxplot indicate the number of individuals.

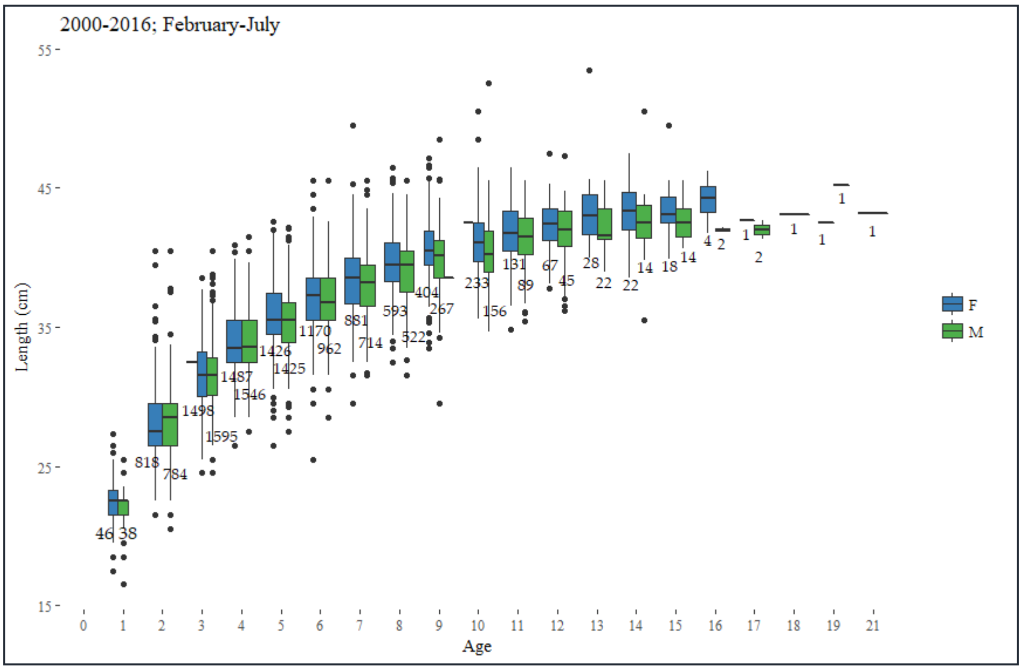

Figure 4.1.3.2. Length distribution by age group and sex for matures (F – females; M – males), from 2000 to 2016 and between February and July. The numbers below each boxplot indicate the number of individuals.

4.1.4 Maturity ogive

The mackerel maturity ogive by sex was reconstructed using the data from 2000 until 2016 (Figure 4.1.4.1). This period was used because the historic data contained a relatively high number of 'old' immatures (see section 4.1.2). The differences observed between sexes, in L50 and K, are probably due to the misclassification errors among immature and spent mackerel at the younger ages, because these are more frequent in females than males (ICES 2015).

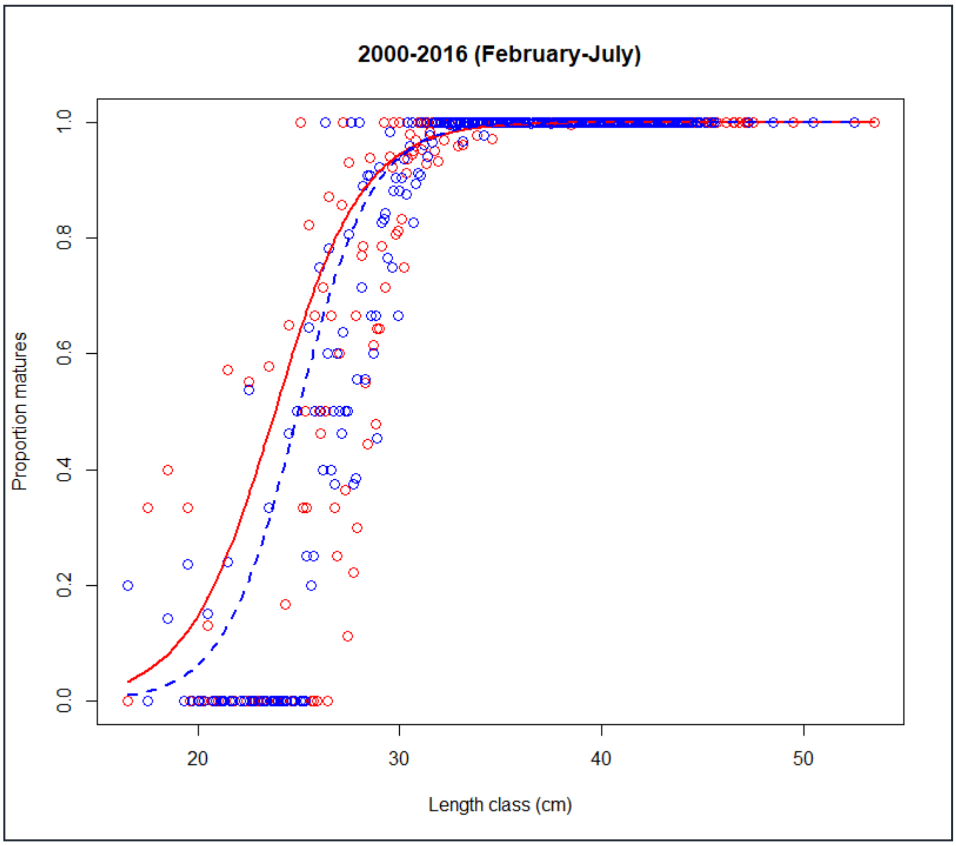

Figure 4.1.4.1. Proportion of matures by length for female (red) and male (blue) mackerel. The dots represent the observations. Female: L50- 23.81; K -0.49; Male: L50- 24.96; K-0.55.

5 Age and maturity staging error matrices (ToR b)

5.1 Age error matrix

Age reading uncertainty can be quantified by an age error matrix (AEM). The elements of an AEM are the probabilities that a sampled fish of true age class *a* is assigned to one of the observed age classes. However, the true age of a fish is usually unknown. A literature review (ICES 2013) showed that several alternatives have been used as "true age" in AEMs. Examples include simulated true age, nearest integer to mean age across readers, modal age, or the age determined using another calcified structure or another method. If an AEM is calculated based on modal or mean age then it will reflect the variance in age readings and not a potential (absolute) bias.

5.1.1 Data on ageing uncertainty

Validated age material is rare. For mackerel, daily ring structures haven been validated in fish larvae (Mendiola and Álvarez 2008). This study may give the potential for validating the first year(s) of growth but, to our knowledge, a validation of the first annual increment(s) based on daily increments has never been carried out. Tagging studies are carried for mackerel (by Norway, Faroe and Iceland) to supplement the data that are used in the assessment. Mark-recapture studies provide a unique basis for age validation studies, but in this case the tagged fish are not actually recaptured. The fish are tagged with radio transmitters and their occurrence is registered by detectors placed on vessels. So, no validated age material is available for mackerel at present. Consequently, no data are available on the accuracy of age determinations, nor can an AEM be calculated based on true ages.

Another source of uncertainty in age reading, referred to as precision, can be estimated by comparing multiple age readings or age readers. Comparison of age readers is frequently done during international age calibration exercises and usually the age determinations of individual age readers are compared to modal age. In the present study we used data from the two most recent calibration exercises for mackerel to calculate AEMs. These exercises consisted of a large-scale otolith exchange conducted in 2008-2009 (Watret *et al*. 2010) and a small-scale otolith exchange conducted in 2014 (Ulleweit 2014). The results of the 2008-2009 exchange were also included in the report of the workshop held in 2010 (WKARMAC, ICES 2010).

Twenty-three readers from 12 countries participated in the 2008-2009 exchange. Fifteen of these readers, from 9 countries, were considered to be experts (Table 5.1.1.1, Watret *et al*. 2010). In the 2014 exchange, 19 readers from 10 countries participated. Only 7 of these readers, from 3 countries, were considered to be experts (Table 5.1.1.1, Ulleweit 2014). Only the age readings of the expert readers were included in the AEM calculations.

The otolith sets of both exchanges contained otoliths from several ICES areas and quarters, but neither sets were truly stratified by area and quarter (Table 5.1.1.2). Furthermore, the areas covered by the 2 exchanges only partly overlapped.

The age range (based on modal age) was similar in both sets (last column of Table 5.1.2.1 and 5.1.2.2), but modal age 10 was missing in the 2008-2009 exchange and modal age 0 in the 2014 exchange. In both exchanges, the age distribution was unbalanced, with too few fish of ages 0 and 9+.

The overall agreement of the expert readers was 68% in the 2008-2009 exchange and 75% in the 2014 exchange. A robust comparison of age reading uncertainty between the 2 exchanges is obstructed by the differences between the (number of) participating readers and by differences in the (spatial, seasonal and age range) coverage of the otolith sets.

To obtain a good estimate of the actual age reading uncertainty in a stock assessment it is important that all age readers who supply data to the assessment participate. Furthermore, the exchange set should cover all ages, seasons and areas included in the assessment adequately.

Table 5.1.1.1. Number of expert and inexperienced age readers by country in each exchange.

Country	Expert readers		Inexperienced readers	
	2008–2009 exc.	2014 exc.	2008–2009 exc.	2014 exc.
Denmark	1			1
Faroe			2	2
France			2	1
Germany	1			1
Iceland			1	1
Ireland	1		1	1
Netherlands	1	1	1	1
Norway	4	3		3
Portugal	1			
Spain	3	3		1
UK-England	2			
UK-Scotland	1		1	
Total	15	7	8	12

Table 5.1.1.2. Number of otoliths by ICES area and quarter in each exchange.

ICES area	2008–2009 exc.				2014 exc	
	Q1	Q2	Q3	Q4	Q1	Q3
2						23
4.a,b	10		20	20	38	37
6.a	20			10	36	
7.b				10	30	
7.j	10					
8.b	15	10				
8.c west		10		10		
8.c east	10			10		
9.a	10	5	15			
Total	75	25	35	60	104	60

5.1.2 Matrix

The AEM based on the age readings of 15 expert readers in the 2008–2009 exchange (Table 5.1.2.1) has been used in the sensitivity analyses (Section 6). This AEM was preferred because of the larger number of participants and the higher coverage of the otolith set. No otoliths with modal age 10 were available. Estimates of the probabilities for age 10 are necessary for the sensitivity analyses. Therefore, probabilities for modal age 10 were obtained by averaging of the probabilities for ages 9 and 11 (diagonally in the matrix).

The AEM based on the age readings of 7 expert readers in the 2014 exchange (Table 5.1.2.2) shows less variance and less skewness compared to the previous AEM, which corresponds with the higher percentage agreement in this exchange. As mentioned before, the differences in the uncertainty estimates between the exchanges may be due to differences in the participants and the coverage of the otolith sets. An exchange should include all readers, ages, seasons and areas included in the assessment to obtain a realistic AEM.

The third AEM (Table 5.1.2.3) illustrates the age reading uncertainty of one (inexperienced) reader compared to the modal age of 7 expert readers. Assuming that the modal age is (close to) the true age then this reader has a clear bias of underestimating the age especially at older ages. Such a skewed AEM will potentially have a larger effect on the outcome of stock assessments than random variation. In the sensitivity analyses (Section 6), the overall catch-at-age matrix is multiplied by the AEM. A more realistic approach might be if country specific AEMs are incorporated at the national level, thus weighing the impact of age reading uncertainty of a certain country by the catch of that country. Such an approach was outside the scope of the present workshop. Moreover, it is doubtful whether this is feasible or worthwhile in future sensitivity analyses.

A probably more important consideration is the true age. All AEMs presented here assume modal age to be a close approximation of true age. However, in the absence of validated (known age) material, it cannot be excluded that a "dissident" reader, for example the reader illustrated in the third AEM, actually gets the ages right and the majority gets it wrong. It is impossible to use only validated material in exchanges, because such material is so rare, but if age readers are calibrated based on validated material then the confidence in the modal age can be much higher.

Table 5.1.2.1. Age error matrix (AEM) of 15 expert readers in the 2008-2009 mackerel exchange. Modal age is assumed to be the true age. The probabilities for modal age 10 are averages of the probabilities for ages 9 and 11.

age	observed													N-fish
modal	0	1	2	3	4	5	6	7	8	9	10	11	12+	
0	0.93	0.05	0	0	0.02	0	0	0	0	0	0	0	0	4
1	0	0.91	0.05	0.04	0	0	0	0	0	0	0	0	0	15
2	0	0.02	0.87	0.09	0.02	0.01	0	0	0	0	0	0	0	25
3	0	0	0.08	0.78	0.11	0.03	0	0	0	0	0	0	0	21
4	0	0	0.01	0.14	0.71	0.11	0.02	0.01	0	0	0	0	0	38
5	0	0	0.01	0.03	0.14	0.67	0.13	0.02	0.01	0	0	0	0	29
6	0	0	0	0.01	0.08	0.23	0.57	0.08	0.02	0.01	0	0	0	22
7	0	0	0	0.01	0.03	0.06	0.18	0.54	0.14	0.05	0.01	0.01	0	13
8	0	0	0	0	0	0.05	0.07	0.20	0.38	0.14	0.08	0.06	0.03	16
9	0	0	0	0	0.01	0.02	0.03	0.12	0.17	0.38	0.16	0.07	0.05	7
10	0	0	0	0.02	0.02	0.04	0.03	0.01	0.09	0.18	0.35	0.20	0.06	0
11	0	0	0	0	0.04	0.04	0.07	0.04	0	0.07	0.18	0.32	0.25	2
12+	0	0	0	0	0	0	0	0.02	0.05	0.07	0.09	0.14	0.64	3

Table 5.1.2.2. Age error matrix (AEM) of 7 expert readers in the 2014 mackerel exchange. Modal age is assumed to be the true age.

age	observed													N-fish
modal	0	1	2	3	4	5	6	7	8	9	10	11	12+	
0														0
1	0	0.98	0.01	0.01	0	0	0	0	0	0	0	0	0	22
2	0	0.02	0.95	0.02	0.01	0	0	0	0	0	0	0	0	33
3	0	0	0.13	0.87	0	0	0	0	0	0	0	0	0	13
4	0	0	0	0.04	0.82	0.09	0.05	0	0	0	0	0	0	7
5	0	0	0	0.02	0.14	0.73	0.09	0.03	0	0	0	0	0	13
6	0	0	0	0	0.01	0.12	0.69	0.13	0.03	0.02	0.01	0	0	19
7	0	0	0	0	0	0.03	0.16	0.67	0.09	0.04	0.02	0.01	0	25
8	0	0	0	0	0	0	0.01	0.17	0.68	0.10	0.02	0.01	0	11
9	0	0	0	0	0	0	0	0	0.04	0.58	0.21	0.17	0	3
10	0	0	0	0	0.01	0	0.01	0.07	0.01	0.15	0.61	0.07	0.06	9
11	0	0	0	0	0	0.03	0	0.03	0.05	0.15	0.15	0.51	0.08	5
12+	0	0	0	0	0	0	0	0	0.03	0	0.09	0.09	0.78	4

Table 5.1.2.3. Age error matrix (AEM) of 1 inexperienced reader in the 2014 mackerel exchange. Modal age of the 7 expert readers is assumed to be the true age.

age	observed													N-fish
modal	0	1	2	3	4	5	6	7	8	9	10	11	12+	
0														0
1	0	1	0	0	0	0	0	0	0	0	0	0	0	22
2	0	0.09	0.91	0	0	0	0	0	0	0	0	0	0	33
3	0	0	0.54	0.46	0	0	0	0	0	0	0	0	0	13
4	0	0	0	0.71	0.29	0	0	0	0	0	0	0	0	7
5	0	0	0.08	0.15	0.46	0.31	0	0	0	0	0	0	0	13
6	0	0	0	0.05	0.47	0.21	0.21	0.05	0	0	0	0	0	19
7	0	0	0	0.08	0.32	0.28	0.24	0.08	0	0	0	0	0	25
8	0	0	0	0	0.09	0.36	0.27	0.09	0.09	0.09	0	0	0	11
9	0	0	0	0	0.33	0	0.33	0	0	0.33	0	0	0	3
10	0	0	0	0	0	0.22	0.22	0.22	0.11	0.00	0.11	0.00	0.11	9
11	0	0	0	0	0	0.20	0.20	0.20	0	0.20	0	0.20	0	5
12+	0	0	0	0	0	0	0.25	0.25	0	0	0	0.50	0	4

5.2 Maturity staging error matrix

The maturity staging uncertainty was quantified, similarly as for age estimation, by an maturity staging error matrix (MSEM). The probabilities that a sampled fish of true maturity class *a* is assigned to one of the observed maturity classes can be assessed more easily for maturity compared to age. "True maturity" can be assessed by histological examination of the gonad. However, assessing true maturity is a time and money consuming exercise and not carried out very often during regular maturity assessment.

5.2.1 Data on maturity staging uncertainty

Two maturity staging workshops have been held for mackerel, first in 2007 (ICES 2007) and the second in 2015 (ICES 2015). The workshop in 2007 focused mainly on preparing an internationally agreed maturity scale. Only a small calibration exercise was carried out to check for difference in maturity staging results between the different scales used. Thus participants examined the same gonads using various scales, which were in use by the participating institutes and the 'new' internationally agreed maturity scale (ICES 2007). The results of the maturity staging differed depending on the maturity scale used (ICES 2007).

The workshop in 2015 focused on calibration between maturity stagers, all using the same internationally agreed maturity scale (ICES 2015), in which data should be reported to the international databases. Two calibration rounds were performed from pictures. The first round used images from Northeast Atlantic and Mediterranean waters throughout the year. Only one immature fish was included in this calibration (Table 5.2.1.1). No histological validation was available for the first round. In the second picture calibration samples were used that were taken during the spawning season and in the spawning area of the western and southern mackerel stock. Four immature fish were included in the second calibration (Table 5.2.1.1). All samples were histologically validated in this second round. Also a third calibration was carried out with frozen mackerel, but these fish were all in maturity stage 4 (ICES 2015).

Table 5.2.1.1. Number of mackerel used per picture staging round in the workshop for sexual maturity staging of mackerel and horse mackerel (WKMSMAC2) in 2015.

Maturity stage	N Fish 1st round	N Fish 2nd round
1	1	4
2	6	11
3	33	31
4	15	17
5	0	0
6	0	0

For the maturity ogive it is essential to include immature fish. The number of immature fish in both maturity staging workshops were low (ICES 2007, 2015). Future maturity staging workshops should take care that samples used in the calibration exercises are stratified over the different maturity stages and include immature fish.

WKMACQI focused on the assessment of NEA mackerel (western, southern and North Sea spawning components). For the maturity staging error matrix to be informative, it should be based on data with the same spatial and temporal coverage as the assessment. For the NEA mackerel assessment only maturity data from the spawning area, (from Portugal up to Iceland) during the spawning season (February–July) is used. The first image calibration round from the 2015 maturity staging workshop included samples from the Mediterranean. Removing these would also remove the one immature fish in the dataset. The second image calibration round used samples from the correct area and timing. The number of immature fish was still low. The frozen fish were taken from the correct area and timing but were all from the same maturity stage.

The MSEM used in the sensitivity analyses is constructed based on the results of the second round of images, as this has the correct spatial and temporal coverage and includes immature fish. For comparison, a MSEM is also constructed on the results of the first image round.

5.2.2 Matrix

The MSEM using histological validated data (Tables 5.2.2.1 & 5.2.2.2) was preferred for the sensitivity analyses (Section 6). These data also had the needed spatial and temporal coverage for this mackerel stock.

The MSEM using the modal maturity stage is shown for comparison (Tables 5.2.2.3 & 5.2.2.4). Results of both MSEM are very similar.

Table 5.2.2.1. Maturity staging error matrix (MSEM) of expert stagers in the 2015 workshop compared to histological stage, using all maturity stages.

	observed					
histology stage	1	2	3	4	5	6
1	0.50	0.25	0	0.25	0	0
2	0.05	0.33	0.59	0.04	0	0
3	0.00	0.11	0.89	0.01	0	0
4	0.29	0.34	0.07	0.30	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0

Table 5.2.2.2. Maturity staging error matrix (MSEM) of expert stagers in the 2015 workshop compared to histological stage, using only immature and mature stage.

	observed	
histology stage	immature	mature
Immature	0.5	0.5
Mature	0.09	0.91

Table 5.2.2.3. Maturity staging error matrix (MSEM) of expert stagers in the 2015 workshop compared to modal stage, using all maturity stages.

	observed					
modal stage	1	2	3	4	5	6
1	0.47	0.33	0	0.2	0	0
2	0	0.58	0.40	0.02	0	0
3	0	0.21	0.71	0.07	0	0
4	0.02	0.24	0.07	0.67	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0

Table 5.2.2.2. Maturity staging error matrix (MSEM) of expert stagers in the 2015 workshop compared to modal stage, using only immature and mature stage.

	observed	
modal stage	immature	mature
Immature	0.47	0.53
Mature	0.01	0.99

6 Sensitivity analyses of mackerel assessment with regards to biological parameter quality indicators (ToR c)

6.1 Methods

Error matrices produced from exchange workshops provide a general idea of the magnitude and statistical distribution of the errors in age or maturity stage determination. However, they cannot realistically be used to correct the historical data, and thereby improve the quality of the data used as input for the assessment. Some types of assessment models explicitly incorporate ageing error distribution (e.g. stock synthesis), but that is not the case for most of the assessment models used at ICES.

To investigate the likely effect of errors in age and maturity stage determination on the mackerel assessment, WKMACQI made the assumption that the data available on the mackerel stock is not affected by any error on age or maturity stage determination, and therefore uses the WGWIDE 2017 data and assessment as a reference (ICES, 2017b). The error matrices are then used to "pollute" these data, and input data for the assessment affected by these errors are generated. The assessment is then run on these data to determine the sensitivity of its output to the errors incorporated.

6.2 Appling ageing and maturity stage errors to the input vectors to the assessment

Ageing errors affect all the input data to the assessment which are structured by age: catch-at-age, age-structured abundance indices, mean fish weight-at-age, proportion mature-at-age. Errors in the determination of maturation stage only affect the proportion of mature fish-at-age.

Ideally, to quantify the impact of these errors in the determination of biological parameters on the different input matrices used in the assessment, one should start from samples with age and maturity stages known without error (true data), apply the error matrices presented in section 4 and then treat this observed data with the same procedures as used in preparation of WGWIDE to obtain the input vectors for the assessment. In practice, however, such "true data" is not available, as all the data available are based on measurements made with error. During WKMACQI, the assumption therefore had to be made that all available datasets (used as the basis for the current assessment) were not affected by errors in the determination of biological parameters, and the error matrices were applied to these data as if it was true data.

During WKMACQI, the procedure used to compute the mean weights-at-age in the stock and proportion of mature fish at age from biological sampling data could be reproduced as the data and scripts used at WGWIDE were available. Ageing errors were applied to the individual fish in the database. This was done by replacing the age value for each fish by a value drawn in [0; 12], with a probability to sample each age equal to the proportions in the age error matrix (Table 5.1.2.3) for the row corresponding to the original age. Once these errors on the age were applied, the normal procedure was used to estimated mean stock weights and proportion of individuals mature at age. The errors on maturity stage determination were applied with the same method.

For the catch-at-age matrix, each country provides their own catch-at-age data, based on quarterly catch data and age-length keys. These national data are then combined by the stock coordinator before WGWIDE. It was impossible to reproduce this procedure as all the necessary data and scripts were not available for WKMACQI. Instead of applying the age errors to the raw data, they were applied to the final product, the whole catch-at-age matrix, which was multiplied by the age error matrix. Similarly, the abundance index-at-age from the IESSNS survey were also multiplied by the ageing error matrix.

The resulting input vectors are presented in Figures 6.2.1-5. Ageing errors tend to smooth the interannual variations in the catches-at-age (Figure 6.2.1). This effect increases with the age of the fish as the accuracy of age reading decreases. The effect on the IESSNS index is similar (Figure 6.2.2). The effect of ageing errors on the stock mean weights is particularly marked for age 1 (Figure 6.2.3), for which the values are higher than the original data. This is likely due to the fact that age 1 fish are typically underrepresented in the samples, and that when applying ageing errors, the number of older (and hence heavier) fish wrongly aged as age 1 fish represent a large proportion of the age 1 fish. This effect is also, though to a much lesser extent, visible for age 2, for which less samples are available than for ages 3-4. The proportions of individuals mature-at-age are affected in a similar way, with slightly higher proportion mature for ages 1-2 when ageing errors are applied (Figure 6.2.4). The maturity staging errors have an even more dramatic impact on the proportions mature-at-age, with substantially higher values for ages 1 and 2 and slightly lower values for older ages (Figure 6.2.5). The error matrix (Table 6.2.2.2) assumes that 50% of the immature fish are wrongly considered mature, and therefore for age 1 and 2, for which the proportion of immature are high (around 90% and 50 % respectively), applying this error matrix increases strongly the proportion mature. For older ages, almost one third of the fish in stage 4 are wrongly considered stage 1 (immature). This explains that the proportion of immature increases for those ages when applying the error matrix.

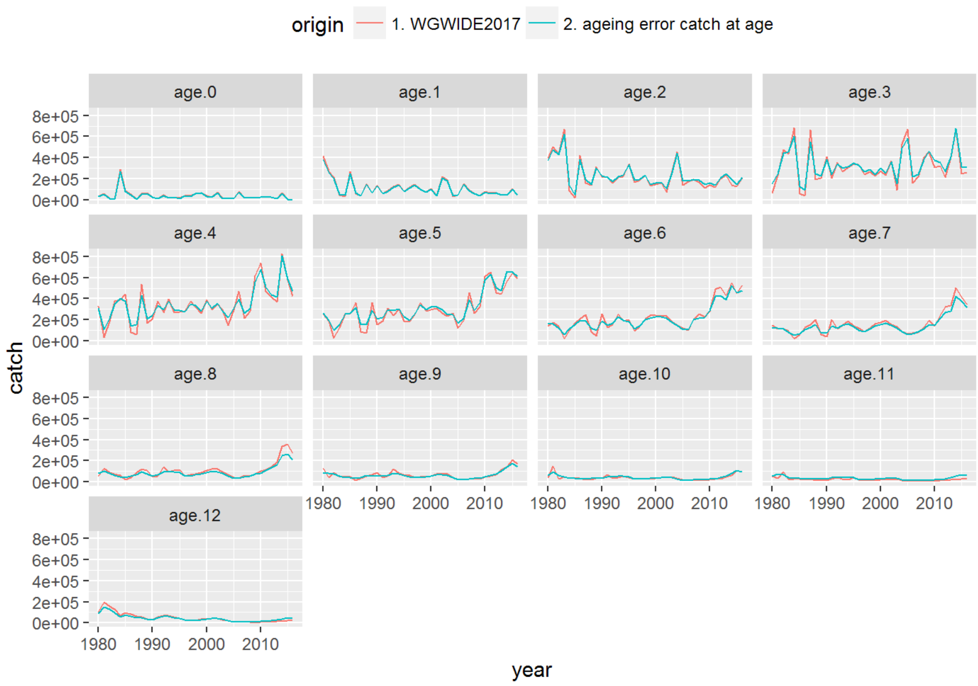

Figure 6.2.1. Comparison of the catch-at-age data used at the 2017 WGWIDE and the catch-at-age affected by ageing errors.

Figure 6.2.2. Comparison of the abundance indices-at-age from the IESSNS used at WGWIDE 2017 and the index affect by ageing errors.

Figure 6.2.3. Comparison of the mean weight-at-age in the stock used at the 2017 WGWIDE and the mean weight-at-age affected by ageing errors.

Figure 6.2.4. Comparison of the proportion of individuals mature-at-age in the stock used at the 2017 WGWIDE and the proportion of individuals mature-at-age affected by ageing errors.

Figure 6.2.5. Comparison of the proportion of individuals mature-at-age in the stock used at the 2017 WGWIDE and the proportion of individuals mature-at-age affected by maturity staging errors.

6.3 Effect on the assessment

6.3.1 Age determination errors

The assessment was run separately with each of the 4 data sources affected by ageing errors, and then once with all 4 data sources.

- Errors on the catch-at-age matrix

Fitting the SAM model on the catch-at-age matrix affected by ageing errors had an effect on the model fit, with small difference in the estimated observation and process standard deviations (Figure 6.3.1.1). As expected, using the catch-at-age matrix affected by ageing errors resulted in a larger observation standard deviation of the catches (i.e. poorer fit). Surprisingly, the standard deviations for most of the other observations (surveys) were also larger for the model fitted using the catch affected by ageing errors. This means that overall, the model provides a poorer fit to the data (except for the recruitment index). The differences observed are however small. The model fitted with the catch-at-age with errors also had smaller standard deviation for the process error and for the recruitment variability. Overall, model uncertainty (on estimates of SSB and F_{bar}) are very similar between the two assessments (not shown).

These small changes in the model parameters had an impact on the states (abundance and fishing mortality-at-age). The estimates for fishing mortality-at-age were different for the older ages since 2000 (Figure 6.3.1.2). Using the catch-at-age data affected by ageing errors resulted in lower F values for ages 6 and 7+ (selectivity plateau). Conversely, starting from the 2000s, the model fitted on the catch-at-age matrix affected by ageing-errors consistently estimated higher abundances for older age groups (7+), up to more than twice larger (for ages 12+) than the model using the true catch-at-age matrix (Figure 6.3.1.3). These differences resulted in marked differences in SSB and F_{bar} (Figure 6.3.1.4).

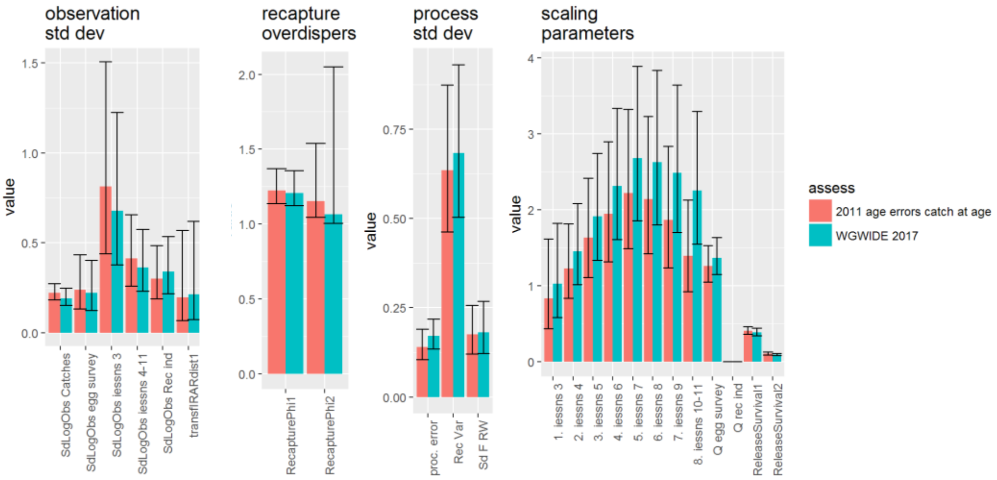

Figure 6.3.1.1. Comparison of the SAM model parameter estimates for the model fitted on the original (WGWIDE2017) catch-at-age matrix and the catch-at-age matrix affected by ageing errors.

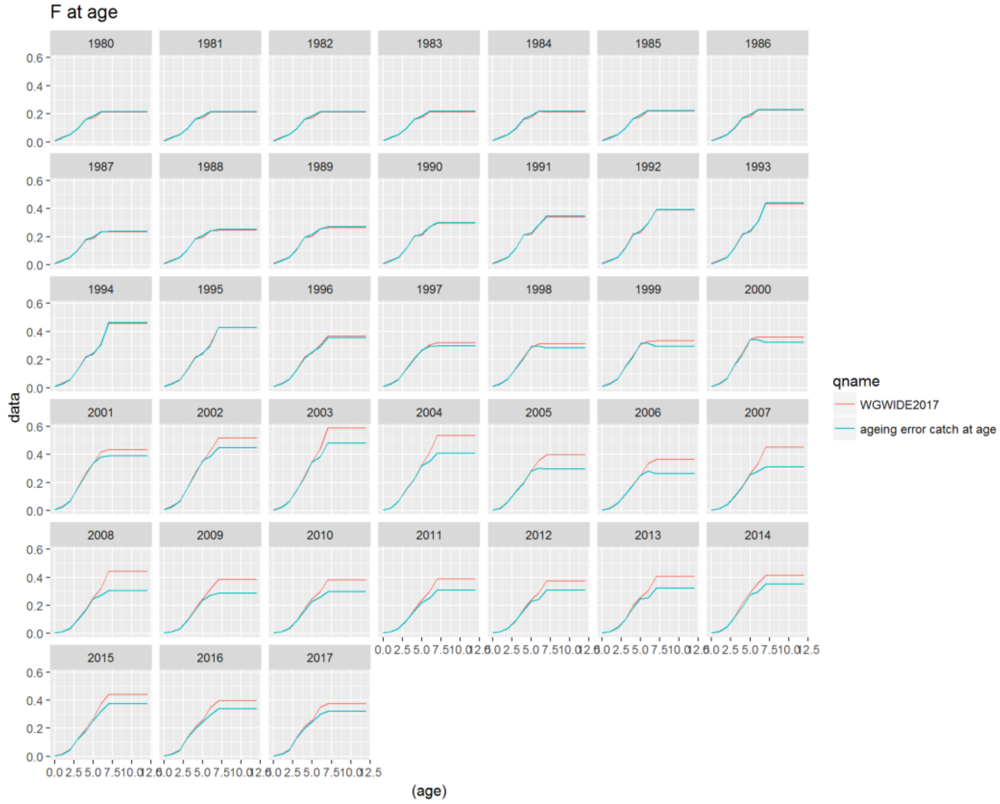

Figure 6.3.1.2. Differences in the estimated fishing mortality-at-age between the mackerel assessment fitted on the original catch-at-age data, and the model fitted on catch-at-age data affected by ageing errors.

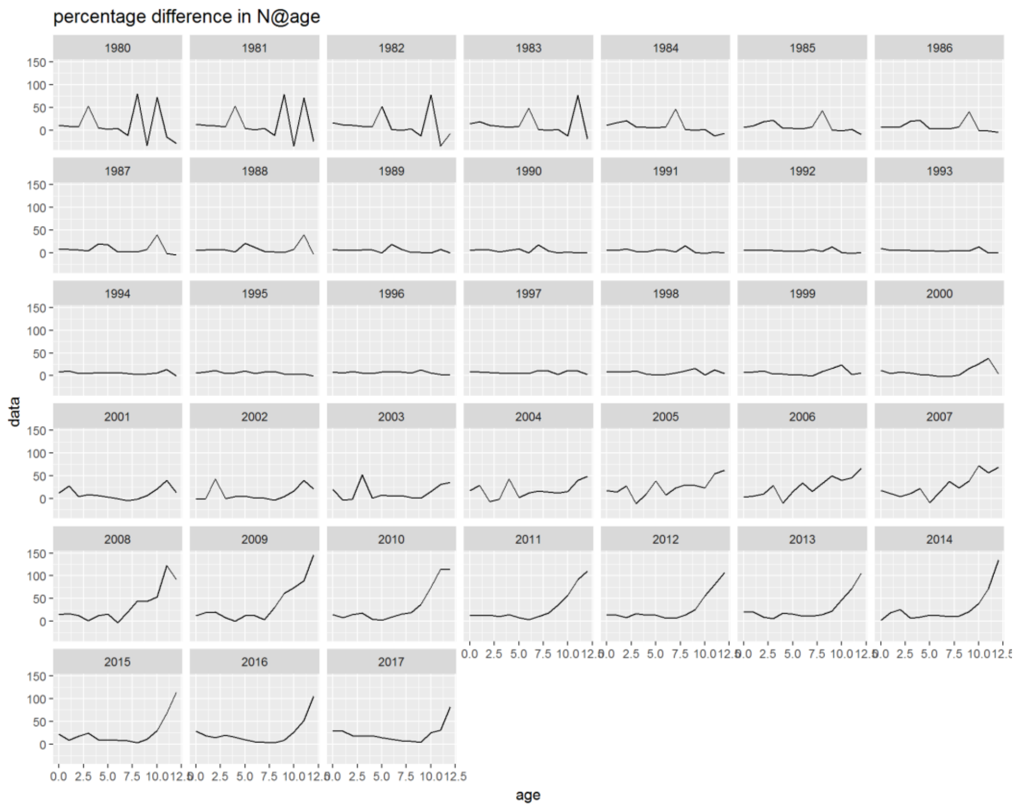

Figure 6.3.1.3. Percentage difference in the estimated abundance-at-age between the mackerel assessment fitted on the original catch-at-age data, and the model fitted on catch-at-age data affected by ageing errors.

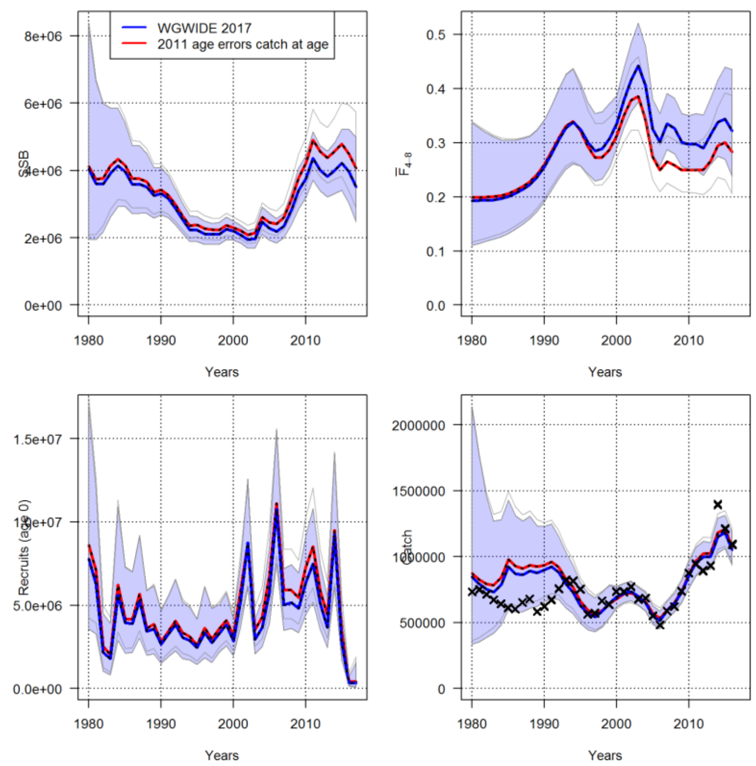

Figure 6.3.1.4. Comparison of the estimated SSB, Fbar, recruitment and modelled catch between the mackerel assessment fitted on the original catch-at-age data, and the model fitted on catch-at-age data affected by ageing errors.

- Errors on the IESSNS index

The model fitted with the IESSNS survey index affected by ageing errors also had slightly different estimated parameters than the original model (Figure 6.3.1.5). Differences were negligible for most parameters, expect for parameters related to the IESSNS survey. The observation standard deviation for this survey was lower for the model fitted with the data affect by ageing errors. This means that, despite the fact that the index is presumably of poorer quality (perturbed by ageing errors), the model shows a better fit to this index than to the original index. The changes in the values of the IESSNS index resulted in changes in the estimated catchability estimates, but with no systematic direction: catchability for ages 3 to 5 and ages 10-11 increased, was unchanged for age 6, and decreased for ages 7 and 8. Model uncertainty was unchanged (not shown).

These changes in the model parameters had an impact on the corresponding states, but the differences were much smaller than for the model with ageing errors on the catch data. Fishing mortality-at-age was virtually unchanged (not shown). Differences were observed for the abundances-at-age in the range of ±20% for the period of years for which the IESSNS index is available (Figure 6.3.1.6). However, there did not appear to be any structure in the observed differences (i.e. to temporal trend or correlation between age groups), indicating that those differences occurred randomly. When calculating age-aggregated quantities, these differences seemed to cancel each other out, as almost no difference was observed in the SSB and Fbar time-series (Figure 6.3.1.7).

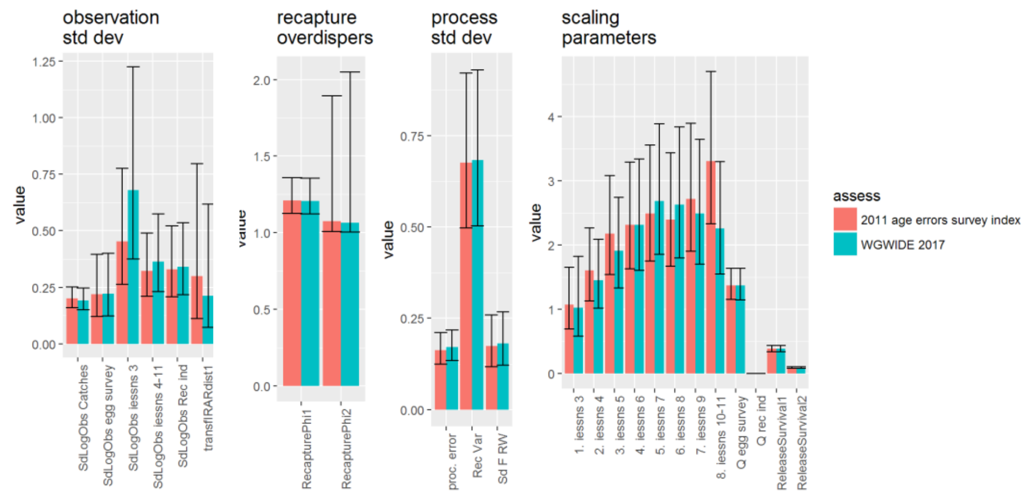

Figure 6.3.1.5. Comparison of the SAM model parameter estimates for the model fitted on the original (WGWIDE2017) IESSNS index and the index affected by ageing errors.

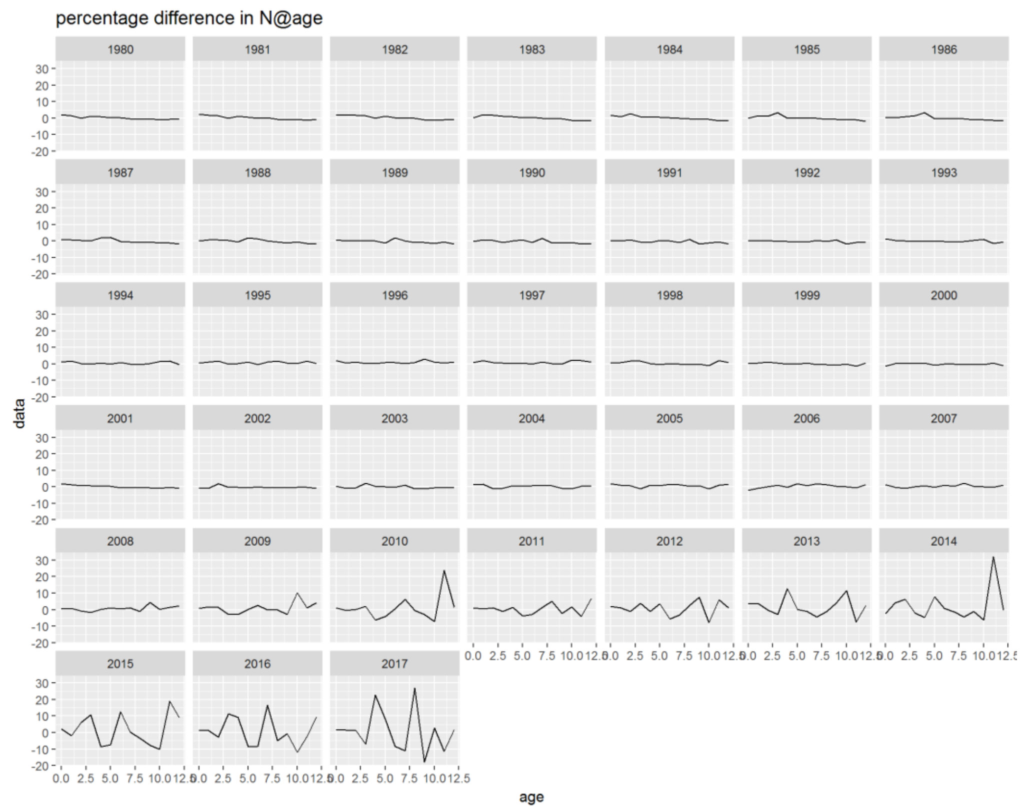

Figure 6.3.1.6. Percentage difference in the estimated abundance-at-age between the mackerel assessment fitted using the original IESSNS index, and the model fitted using the index affected by ageing errors.

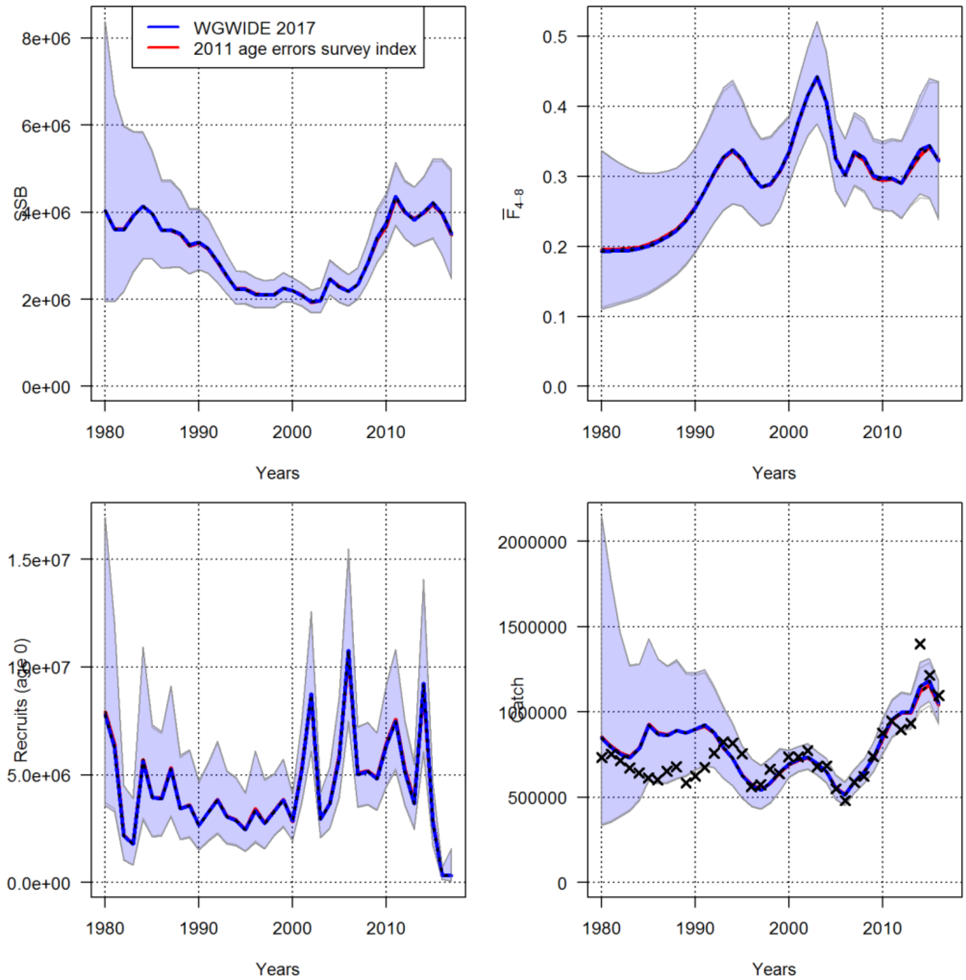

Figure 6.3.1.7. Comparison of the estimated SSB, F_{bar}, recruitment and modelled catch between the mackerel assessment fitted using the original IESSNS index, and the model fitted using the index affected by ageing errors.

- Errors on mean weights-at-age in the stock

In the model fitting process, this input data are used for the computation of the modelled SSB, which is then used for the fitting to the egg survey index. Implementing the effect of ageing errors on the stock weights had almost no effect on the parameters estimated (Figure 6.3.1.8). In particular, the fit to the observations was unchanged. There was only a minimal difference in the estimated catchability for the egg survey.

Given that model parameters were almost identical, not difference was observed in the abundances and fishing mortality-at-age estimates (maximum 0.4% difference for abundances, not shown). Given that the mean weights in the stock are included in the calculation of the SSB, a small difference was observed in the SSB (Figure 6.3.1.9), which also explained the slightly different catchability for the egg survey.

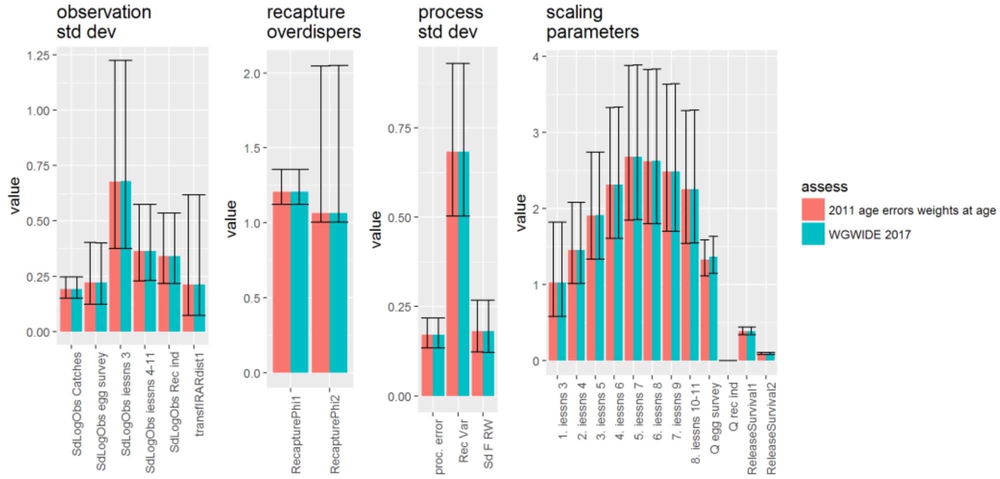

Figure 6.3.1.8. Comparison of the SAM model parameter estimates for the model fitted on the original (WGWIDE2017) stock mean weight-at-age matrix and the stock mean weight-at-age matrix affected by ageing errors.

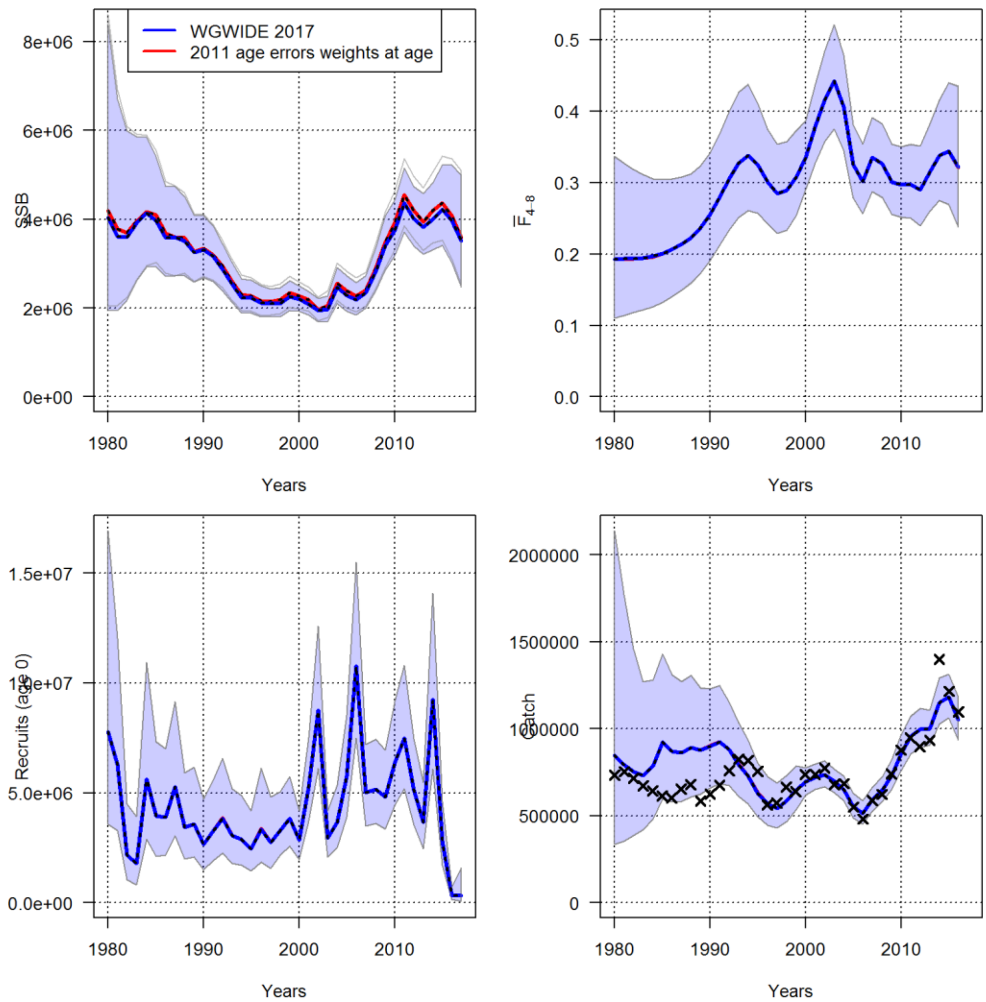

Figure 6.3.1.9. Comparison of the estimated SSB, F_{bar}, recruitment and modelled catch between the mackerel assessment fitted on the original weight-at-age data, and the model fitted on weight-at-age data affected by ageing errors.

- Errors on proportion individuals mature-at-age

As for mean weights-at-age in the stock, the proportion fish mature-at-age only influences model fit through the calculation of the SSB from the estimates states, and the consequences thereof for the fit to the egg survey. In this case, almost no difference was observed neither in the parameters nor in the states.

- Effect of ageing errors on all input data combined

The parameters estimated when the model is fitted with the 4 input data affected by ageing errors simultaneously were nearly identical to the parameters obtained when ageing errors affected the catch-at-age matrix only (Figure 6.3.1.1). The differences observed in the states were also similar to those observed for the model with errors on the catches only (Figure 6.3.1.3).

The effect on the estimated SSB of ageing error on the catch-at-age matrix and on the stock weights seemed to add up, as the difference observed for the model with all data sources affected by ageing errors is roughly the sum of the differences observed when each of these 2 data sources were affected (Figure 6.3.1.10).

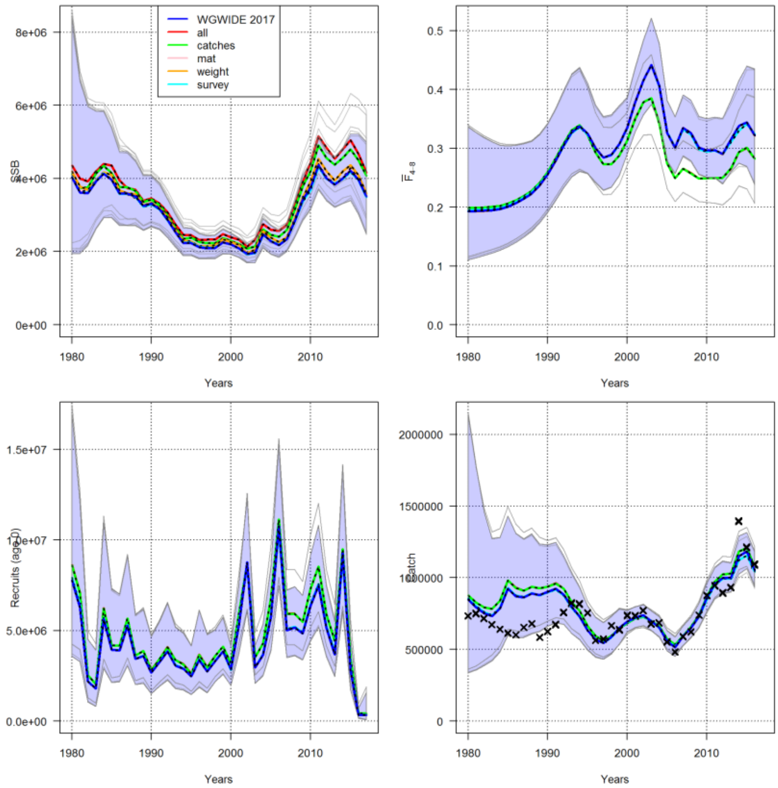

Figure 6.3.1.10. Comparison of the estimated SSB, F_{bar}, recruitment and modelled catch between the mackerel assessment fitted on the original data, and the assessment fitted on each of the data sources, separately and simultaneously, affected by ageing errors.

6.3.2 Maturity stage determination errors

As it was the case for ageing errors, using proportion fish mature-at-age affected by errors in maturity stage determination had only a negligible effect on the estimated parameters and states (not shown).

The estimated SSB showed some small differences with the base case model for some periods of years (e.g. early 2000s, mid 2000s, 2014–2015; Figure 6.3.2.1). The estimates of F_{bar} are nearly identical).

Given the substantially higher proportion fish mature at age 1 (and to some extent 2; Figure 6.2.5) in the data affected by staging errors, the low sensitivity of the assessment may seem surprising. Taking the example of the year 2016, the spawning biomass corresponding to age 1 fish indeed increased substantially, but still represents overall a small part of the SSB, due to the low proportion of fish mature and the low individual weight of the fish at age 1 (Figure 6.3.2.2). Therefore, the higher proportion of age 1 fish mature with staging error has little impact on the SSB. The SSB fraction corresponding to age 2 is larger, and the spawning biomass of age 2 fish also increased when the data affected by maturity staging error is used. However, for older age classes (3 to 6), representing the bulk of the SSB, the lower proportion of fish mature with the staging errors led to a smaller spawning biomass for these age classes. Therefore, as they affect mainly an age group that do not contribute substantially to the SSB, and as the opposite effect, though weaker, is observed for age groups that contribute to most of the SSB, the differences in proportion mature at age do not overall modify the SSB.

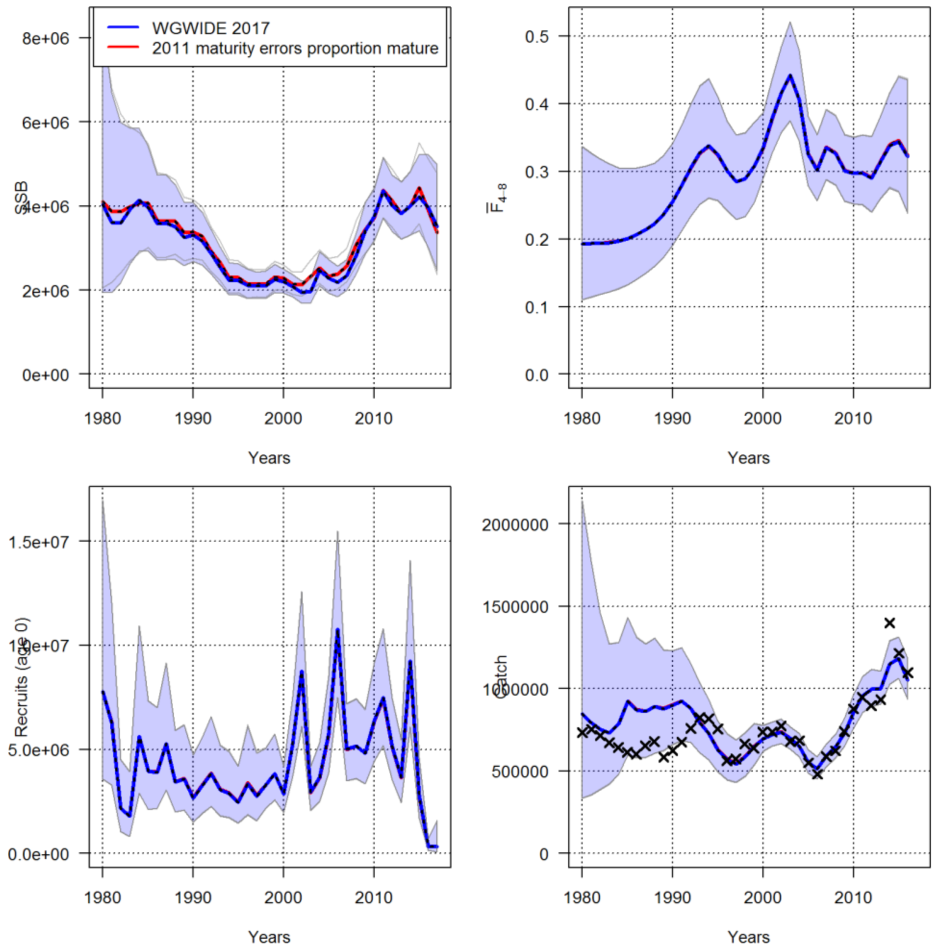

Figure 6.3.2.1. Comparison of the estimated SSB, F_{bar}, recruitment and modelled catch between the mackerel assessment fitted on the original proportion mature-at-age data, and the model fitted on proportion mature-at-age data affected by maturity staging errors.

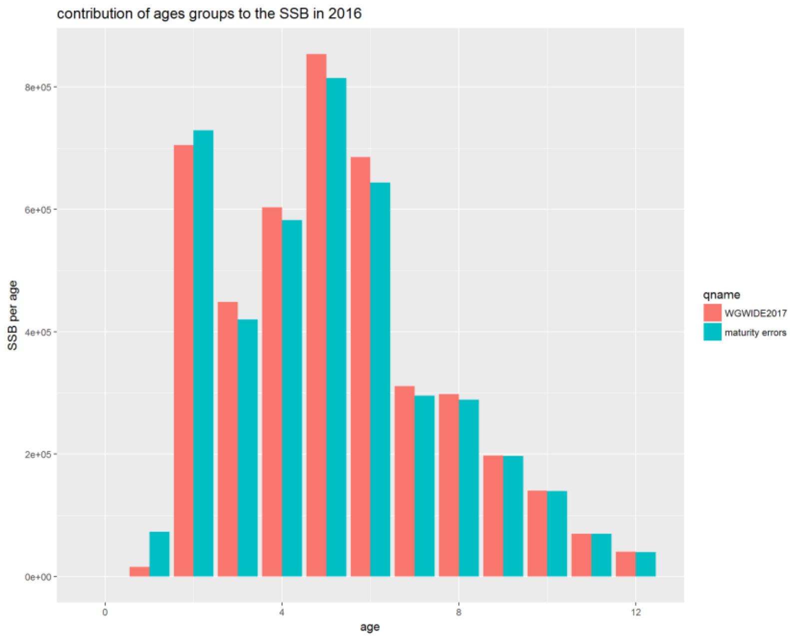

Figure 6.3.2.2. Comparison of age composition of the 2016 SSB between the mackerel assessment fitted on the original proportion mature-at-age data, and the model fitted on proportion mature-at-age data affected by maturity staging errors.

7 Discussion

Biological parameters

Biological parameters used in the mackerel assessment are length, weight, age, sex, maturity, natural mortality, stock structure and indirectly egg stage and fecundity (for the estimation of the egg survey index). Only for age and maturity, data were available to quantitatively assess the quality of these parameters and carry out a sensitivity analyses of the mackerel assessment. The quality indicator data for age and maturity come from international age reading and maturity workshops. Workshops are also carried out for egg identification and staging and fecundity estimation. Data from these workshops could be used to estimate quality of the estimation of these parameters. This uncertainty should be used in the estimation of the index of the egg surveys.

To obtain a good estimate of the actual age reading and maturity staging uncertainty, it is important that all age readers and maturity stagers who supply data to the assessment participate in the workshops and exchanges. Furthermore, the sample set used should cover all ages, maturity stages, seasons and areas included in the assessment adequately.

Specific recommendations for ageing and maturity staging workshops and exchanges are:

- For age calibration exercises, samples should be stratified by age. This is already mentioned in the WGBIOP Guidelines for Otolith Exchanges and Workshops on Age Reading Calibration but should be more emphasised, as it was not taken into account for the mackerel exchanges and workshops.
- To create an AEM representative for stock assessment, the age range, spatial and seasonal coverage of the age calibration exercise should correspond with those used in the assessment.
- For maturity staging workshops samples should be stratified by maturity stage. It is important that the samples contain sufficient immatures, as maturity ogives (immature vs. mature) are used in assessments. The WGBIOP guidelines for Workshops on Maturity Staging should be updated with this recommendation.
- For maturity calibrations, the age of the fish sampled should also recorded, because the maturity ogive used in assessments is age-based. This needs to be added to the WGBIOP guidelines for Workshops on Maturity Staging.
- Validation of the maturity stage is 'relatively' easy, with histological examination and should be carried out in maturity staging workshops.
- Participation in ageing and maturity staging calibrations should at least be from all who deliver data for the assessment. To get a good estimate of uncertainty, all who deliver data should be involved. Synchronisation of ageing and maturity staging calibrations with benchmarks might improve involvement of those who deliver data for assessments.
- Often maturity staging exercises have a problem with lack of material available for the workshops. WGBIOP should recommend to all survey groups and WGCATCH to take a few extra samples for calibration exercises during their regular sampling, i.e. some samples each survey. This will spread the burden of collection of samples among all participants and ensure that samples are collected over the entire spatial and temporal coverage used for the assessment.

- In order to enhance chances of getting fresh samples, maturity staging workshop should be conducted during the main spawning period.
- At least one of the samples used in the maturity staging workshops should be fresh, in order to allow performing the microscopic validation of the maturity stages.

Model sensitivity

The result of the analyses presented here indicate that errors in biological parameter determination can have a substantial effect on the output of the mackerel assessment. For instance, when ageing errors are affecting all data sources, a difference of +14% in the SSB and –14% in F_{bar} is observed in the last 5 years. This is a substantial difference, and the SSB and F_{bar} trends of the model based on data affected by ageing errors are close to the limit of the confidence bounds of the WGWIDE2017 assessment. Most of this difference is explained by the impact of the errors on the catch-at-age matrix (+13.5% in SSB and -14% in F_{bar}). This result contrasts with the outcome of a similar study conducted in 2011 (Brunel, 2011), using ageing errors from the same workshops (ICES, 2010), which found a maximum discrepancy of ±5% in the SSB. The 2011 assessment was based on ICA, using only the catches and egg survey index. The model had no process deviations, and was not free to estimate the magnitude of observation errors (and arbitrary weighting was used). In the former ICA assessment, errors in the catch-at-age matrix automatically resulted in errors of the same sign in the estimated numbers-at-age. Since positive and negative errors in the catch-at-age matrix tend to alternate from one cohort to the other, the errors on the abundance-at-age cancelled each other out and the impact on the SSB was small. In the case of the SAM mackerel assessment, using catch-at-age data with error affects the estimated parameters, i.e. gives a different weight to the different data sources, which seems to have larger consequences on the estimated abundances-at-age than in a simpler model as ICA. This highlights the fact that different types of models may be affected in different ways by the errors in the input data related to errors in biological parameters determination. The sensitivity of assessment methods to these errors should be investigated in a more systematic way to understand the model specific consequences of these errors. Such an analysis could be based on simulated populations, from which assessment input could be derived, with varying degrees of accuracy.

The analyses done here show that errors in the determination of biological parameters affected the mackerel assessment at different levels:

- Impact of errors that affect the observations (catches and survey indices), have a high likelihood of resulting in a different model fit (different observation and process variances). This effect on model fit also implies different estimated abundances and fishing mortality-at-age. This was observed here for the effect on catch-at-age and for the IESSNS index. The sensitivity to errors in the catch-at-age matrix was larger than the sensitivity to errors in the IESSNS. The Model gives a larger weight to the catch information compared to the IESSNS survey. Errors affecting the catch data have therefore more impact than errors affecting the survey index.
- Impact of errors on input data other than observations (stock weights and proportion fish mature) had little impact on the model fit. These input data enter in the calculation of the SSB (from the abundance and fishing mortality-at-age estimates), and errors affecting this input data will likely have an impact on the SSB. This was the case here for the effect of errors on stocks

weight, but not on maturity-at-age, for which the effect on different age classes cancelled each other out.

- Since the modelled SSB is impacted, this could potentially have consequences for model fit as well (through the fit to the egg survey index). However, in the analyses presented here, the difference in SSB was not large enough to substantially affect model parameters, and the estimated abundance-at-age in the stock were not affected.

Usability for other species and assessment models

It should however be noticed that the sensitivity to errors on biological parameter determination would likely be different for assessment of other stocks, even based on SAM, regardless of the stock specific magnitude of the ageing or maturity staging errors. The mackerel assessment is particularly unstable (i.e. small changes in the data, or addition of new data can result in substantially different perception of the stock). The magnitude of the differences observed here in the different runs is in par with the magnitude of the yearly revisions of stock trends having occurred in the recent years. However, SAM assessments based on more data, or on data of better quality, may not be affected in a similar manner as the mackerel assessment. Given the flexibility of the SAM model, particularly with respect to how the variability of the input data are handled (i.e. being ascribed to observation or to process deviations), it is difficult to make any prediction of how other SAM assessments may react to ageing or maturity staging errors.

Cooperation between assessment and biological experts

The workshop was an excellent opportunity to bring together people from the assessment and biology side to actually work together on biological data. This increased the understanding of the uncertainties in biological data, how biological data are used in assessments, and what data are needed to evaluate the effect of uncertainties on the outcome of the assessment. Such a close collaboration cannot be achieved by either assessors participating in WGBIOP or survey people participating in WGWIDE. Both at WGWIDE and WGBIOP the focus is too much on the work at hand and there is not sufficient time to scrutinise the data and the assessment results together.

7.1 Follow-up work

The analyses of the available maturity data for mackerel showed that some older fish were noted as immature (Chapter 4), also the maturity ogive shows outliers, especially for females (Figure 4.1.4.1). It is impossible to validate the historic maturity data, as the samples are not stored. But a relatively easy exercise could be carried out to check this problem. Collecting age and gonads of immature mackerel could be carried out during the mackerel and horse mackerel egg survey in 2019. Macroscopic and microscopic maturity staging of these samples, together with the regularly collected and analysed fecundity samples will allow for an update of the maturity ogive for mackerel. With these data the MSEM can also be checked.

Maturity staging can be relatively easily validated with true maturity, for ageing this is much more difficult, because validated material is scarce. All AEMs presented in this report assume modal age to be the true age. It can however not be excluded that a "dissident" reader, actually gets the ages right and the majority gets it wrong. Table 5.1.2.3 indicates the bias if this were the case. It needs to be checked what the effect would be on the assessment of such a scenario.

The SAM assessment model is flexible and the mackerel assessment is particularly unstable. The results may be different if the same AEM and MSEM were applied to other species and other assessment models. The sensitivity of assessment methods to these errors should be investigated in a more systematic way to understand the model specific consequences of these errors. Such an analysis could be based on simulated populations, from which assessment input could be derived, with varying degrees of accuracy.

8 References

Brunel, T. 2011. Influence of ageing errors on the NEA Mackerel assessment outputs. Woking Document presented to the 2011 WGWIDE meeting.

Gonçalves, P., Ávila de Melo, A., Murta, A. G., Cabral, H. N. 2017. Blue whiting (*Micromesistius poutassou*) sex ratio, size distribution and condition patterns off Portugal. Aquatic Living Resources, 30, 24.

ICES, 2007. Report of the Workshop on Sexual Maturity Staging of Mackerel and Horse mackerel (WKMSMAC). ICESCM2007/ACFM:26, 52 pp.

ICES, 2010. Report of the workshop on age reading of mackerel (WKARMAC), 1-4 November 2010, Lowestoft, UK. ICES CM 2010/ACOM: 46.

ICES, 2013. Report of the Second Workshop of National Age Readings Coordinators (WKNARC2), 13-17 May 2013, Horta, Azores. ICES CM 2013/ACOM:52

ICES, 2015. Report of the Workshop on Maturity Staging of Mackerel and Horse Mackerel (WKMSMAC2). ICES CM 2015/SSGIEOM:17, 93 pp.

ICES, 2016. Report of the Working Group on Biological Parameters (WGBIOP). ICES CM 2016/SSGIEOM:08, 109 pp.

ICES, 2017a. Report of the Working Group on Biological Parameters (WGBIOP). ICES CM 2017/SSGIEOM:08, 132 pp.

ICES, 2017b. Report of the Working Group on Widely Distributed Stocks (WGWIDE). ICES CM 2017/ACOM:23, 1012 pp.

ICES, 2017c. Final Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys. ICESCM 2017/SSGIEOM:18, 136 pp.

ICES, 2018. Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys (WGMEGS). ICESCM 2018/EOSG: 17, 74 pp.

Mendiola, D. and Álvarez P., 2018. Validation of daily increments in the otolith microstructure of Northeast Atlantic mackerel fish larvae. Fisheries Research 89: 300-304.

Ulleweit, J., 2014. Small scale otolith exchange for North East Atlantic Mackerel (*Scomber scombrus*) 2014. Thünen Institute of Fisheries Internal Report.

Watret, R., Goudie, O. Shanks, A., 2010. Analysis of data from mackerel (*Scomber scombrus*) otolith exchange scheme 2008/2009. Marine Scotland Science Internal Report No 02/10.

Annex 1: List of participants

NAME	INSTITUTE	COUNTRY (OF INSTITUTE)	E-MAIL
Loes Bolle	Wageningen Marine Research P.O. Box 68, 1970 AB IJmuiden, The Netherlands	The Netherlands	loes.bolle@wur.nl
Thomas Brunel	Wageningen Marine Research P.O. Box 68, 1970 AB IJmuiden, The Netherlands	The Netherlands	thomas.brunel@wur.nl
Liz Clarke (by corespondence)	Marine Scotland Science Marine Laboratory 375 Victoria Road Aberdeen AB11 9DB UK	UK, Scotland	L.Clarke@MARLAB.AC.UK
Cindy van Damme (chair)	Wageningen Marine Research P.O. Box 68, 1970 AB IJmuiden, The Netherlands	The Netherlands	cindy.vandamme@wur.nl
Patricia Gonçalves	Portuguese Institute for the Sea and the Atmosphere (IPMA) Avenida de Brasilia 1449 – 006 Lisbon Portugal	Portugal	patricia@ipma.pt
Pablo Carrera López (by correspondence)	Instituto Español de Oceanografia Calle del Corazón de María 8, 28002 Madrid, Spain	Spain	pablo.carrera@ieo.es
Begoña Villamor (by correspondence)	Instituto Español de Oceanografia Calle del Corazón de María 8, 28002 Madrid, Spain	Spain	begona.villamor@ieo.es

Annex 2: Agenda

Tuesday 15 May

9:00 Welcome and introductions

9:30 Background, introduction and setup of WKMACQI

10:15 Presentation on Age Error Matrix (Loes)

13:30 Presentation on Age Error Matrix sensitivity analyses (Thomas)

10:45 ToR a: Review and consider quality indicators for and issues with biological parameters of western, southern and North Sea mackerel. Please come prepared with information for this ToR

13:00 ToR b: Age Error Matrix and Maturity Staging Error Matrix

Wednesday 16 May

9:00 Report back on ToR a and progress ToR b

9:30 ToR c: Carry out sensitivity analyses of the mackerel assessment with regards to the quality indicators of mackerel biological parameters.

13:00 Setup reporting/manuscript

13:30 Continue work on various ToR's and report writing

Thursday 17 May

9:00 Report back on work carried out on various ToR's

9:30 Recommendations, follow-up of the WK, etc.

10:15 Continue work on various ToR's and report writing

13:00 Continue work on various ToR's and report writing

15:15 Final plenary on recommendations, follow-up of the WK, report sections

17:00 End of workshop

Annex 3: Recommendations

Recommendation	Adressed to
1. To optimise the estimation of error matrices WKMACQI recommends (see also Chapter 7): Follow the WGBIOP guidelines for ageing workshops and exchanges, and update and emphasise the guidelines to stratify samples for calibrations by age. Stratify samples for maturity staging workshops by maturity stage (including immature fish). The WGBIOP guidelines should be updated with this recommendation. Follow the WGBIOP guidelines and include a validation by histology in maturity staging workshops. Include length and age of the fish in the reporting of the results of the maturity staging workshops. Certify that the age range in age calibration exercises corresponds with the age range used in the assessment. Certify that spatial and temporal coverage of the samples used on both age and maturity calibration exercises correspond with the coverage in the assessment. Confirm that all who deliver data for the assessment participate in age and maturity calibration exercises. Synchronisation of workshops and exchanges with the benchmarks of stocks might improve involvement in the calibrations. Not necessary for the creation of error matrices but to improve maturity staging workshops, WKMACQI recommends to: Preferentially conduct the maturity staging workshop during the main spawning period. This will considerably enhance the chance of getting fresh samples in the most reliable period for macroscopic maturity staging. At least one of the samples used in the maturity staging workshops should be fresh, in order to allow performing the microscopic validation of the maturity stages.	WGBIOP
2. During regular samplings a few extra samples should be collected for future ageing and maturity staging exercises. This ensures enough samples from the correct areas are available for future calibrations and the collection of samples is spread among all participating countries. WGBIOP should prepare general guidelines to be included in sampling manuals. WGBIOP should disseminate this recommendation to WGBIFS, WGMEGS, WGACEGG, WKFATHOM, PGDATA, WGIDEEPS, WGNEACS, WGBEAM, WGCATCH, IBTSWG and WGIPS	WGBIOP
3. WKMACQI recommends to estimate the error in egg identification and staging for the mackerel egg surveys and check the effect of this error on the SSB index from the egg survey and prepare advice for the assessment with regards to this error.	WGMEGS, WKFATHOM, WGBIOP
4. During the 2019 mackerel egg survey, mackerel gonads from immature fish by age should be collected for histology and be used to construct a microscopic maturity ogive for stock assessment. (See also Chapter 6.1).	WKFATHOM, WGMEGS, WGACEGG,