SmartDots Summary for event 3985

1 Executive summary

This report presents the outcomes of the advanced readers of SmartDots Event 3985, an international age-reading intercalibration exercise for megrim (*Lepidorhombus whiffiagonis*) within ICES areas 27.7b–k and 8abd. This exercise was conducted as a preliminary intercalibration ahead of the WKARMEG workshop to assess the consistency of age determination among participants. It also aimed to identify the main sources of disagreement and the most challenging aspects of age determination, thereby highlighting areas requiring further discussion during the workshop. The ultimate objective was to reduce uncertainty in the biological inputs used for stock assessments.

A total of 101 whole otoliths were analyzed by advanced readers (reading for stock assessments). Age estimates were evaluated using standard precision metrics, including percentage agreement (PA), coefficient of variation (CV), average percentage error (APE), and age bias analyses. In addition, a multistage modal age estimation approach, incorporating reader expertise, was applied alongside the traditional unweighted method.

Results indicate moderate overall agreement among advanced readers (PA 64%, CV 17%, APE 11%), with variability concentrated in younger and older age classes. The weighted mean relative bias among advanced readers was close to zero, indicating the absence of a strong systematic ageing bias at the group level. The traditional modal age approach produced multiple modal ages in 18% of cases, whereas the multistage approach successfully resolved all ambiguities, demonstrating its effectiveness in improving consistency of assigned ages. The main issues identified were the recognition of the first annulus, the identification of split rings, and the interpretation of the otolith edge.

Overall, the exercise confirmed that age estimation for this species remains subject to reader-dependent variability, particularly at the extremes of the age range. It is recommended for WKARMEG to focus on standardizing interpretation criteria, addressing difficult age classes, and strengthening training to further reduce inter-reader variability and improve the robustness of age data used in ICES assessments. The follow-up workshop WKARMEG is foreseen in June 2026 in Boulogne to resolve the issues.

2 Overview of samples and advanced readers

Table 1: Overview of samples used for the exchange event3985. The modal age range for all samples is 1-10.

Year	ICES area	Strata	Quarter	Number of samples	Modal age range	Length range
2021	27.7.e	Strata_WH	3	1	6	375 mm
2022	27.7.e	Strata_WH	4	1	6	405 mm
2022	27.8.a	Strata_WH	2	4	2-7	180-395 mm
2023	27.7.g	Strata_WH	2	3	4-10	295-490 mm
2023	27.8.a	Strata_WH	3	1	8	275 mm
2024	27.7.b	Strata_WH	3	2	1-3	95-160 mm
2024	27.7.c.2	Strata_WH	3	2	4-10	155-480 mm
2024	27.7.e	Strata_WH	2	3	2-8	175-455 mm
2024	27.7.e	Strata_WH	3	2	1-5	180-285 mm
2024	27.7.f	Strata_WH	2	4	3-7	230-410 mm
2024	27.7.f	Strata_WH	3	3	3-8	270-480 mm
2024	27.7.f	Strata_WH	4	2	6-7	300-400 mm
2024	27.7.g	Strata_WH	4	1	8	350 mm
2024	27.7.h	Strata_WH	2	1	5	270 mm
2024	27.7.h	Strata_WH	3	3	5-9	285-475 mm
2024	27.8.a	Strata_WH	1	5	2-8	155-410 mm
2025	27.7.b	Strata_WH	3	3	2-4	135-200 mm
2025	27.7.c.2	Strata_WH	3	2	2-9	130-555 mm
2025	27.7.c.2	Strata_WH	4	9	2-10	120-335 mm
2025	27.7.g	Strata_WH	1	1	4	280 mm
2025	27.7.h	Strata_WH	1	3	5-8	285-440 mm
2025	27.7.j	Strata_WH	1	3	7-10	415-530 mm
2025	27.7.j	Strata_WH	2	8	6-10	310-530 mm
2025	27.7.k.2	Strata_WH	2	19	3-10	260-420 mm
2025	27.7.k.2	Strata_WH	3	4	1	145-150 mm
2025	27.7.k.2	Strata_WH	4	1	9	270 mm
2025	27.8.b	Strata_WH	1	10	2-9	210-445 mm

Table 2: Overview of number of readers.

Expertise	strata	N_readers
Advanced	Strata_WH	6

3 Results overview

3.1 Multimodal cases

Table 3: Summary of statistics; Total number of samples (NSample), a percentage of cases (fish samples) with multiple modes depending on the approach to weight the experience of the reader which will be considered when defining the fish age mode. PercMM_traditional shows the percentage of the total samples for which multiple modes are obtained when all the readers are equally weighted. PercMM_linear_weight shows the percentage of the total samples for which multiple modes are obtained when the weight assigned to the different readers decreases linearly with the experience, while in the PercMM_negexp the weight applied decreases with a negative exponential shape with the experience. The PercMM_multistage shows the percentage of multiple mode cases when a combination of the different methodologies is used, as explained in the material and methods section

NSample	PercMM_traditional	PercMM_linear_weight	PercMM_negexp_weight	PercMM_multistage
101	18 %	2 %	0 %	0 %

3.2 Age readings

Table 4: Age reading table shows the number of readings by modal age.

Modal age	total
1	**24**
2	**72**
3	**42**
4	**78**
5	**66**
6	**74**
7	**48**
8	**78**
9	**59**
10	**54**
11	**6**
Total	**601**

3.3 CV table

Table 5: Coefficient of Variation (CV) table presents the CV per modal age for all advanced readers combined.

Modal age	all
1	**61 %**
2	**29 %**
3	**21 %**
4	**20 %**
5	**9 %**
6	**11 %**
7	**10 %**
8	**9 %**
9	**15 %**
10	**12 %**
11	**15 %**
Weighted Mean	**17 %**

3.4 PA table

Table 6: Percentage agreement (PA) table represents the PA per modal age for all advanced readers combined.

Modal age	total
1	**50 %**
2	**68 %**
3	**69 %**
4	**73 %**
5	**79 %**
6	**69 %**
7	**62 %**
8	**64 %**
9	**46 %**
10	**48 %**
11	**33 %**
Weighted Mean	**64 %**

3.5 APE table

Table 7: Average Percentage Error (APE) table represents the APE per modal age for all advanced readers combined.

Modal age	all
1	**49 %**
2	**20 %**
3	**11 %**
4	**11 %**
5	**5 %**
6	**7 %**
7	**6 %**
8	**5 %**
9	**11 %**
10	**9 %**
11	**13 %**
Weighted Mean	**11 %**

3.6 Relative bias table

Table 8: The relative bias (as the difference between the mean and modal age) per modal age for all advanced readers combined.

Modal age	all
1	**0.96**
2	**0.12**
3	**0.00**
4	**-0.06**
5	**0.03**
6	**0.13**
7	**0.00**
8	**0.05**
9	**-0.33**
10	**-0.28**
11	**-1.00**
Weighted Mean	**0.00**

3.7 Bias plot

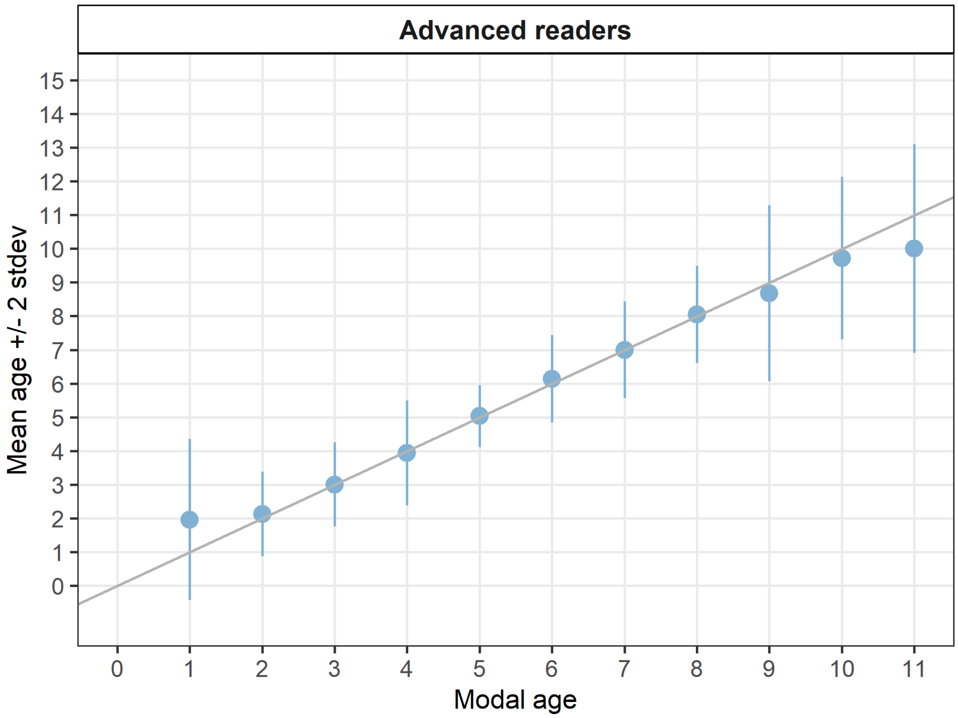

Figure 1: Age bias plot for advanced readers.

3.8 Growth analysis

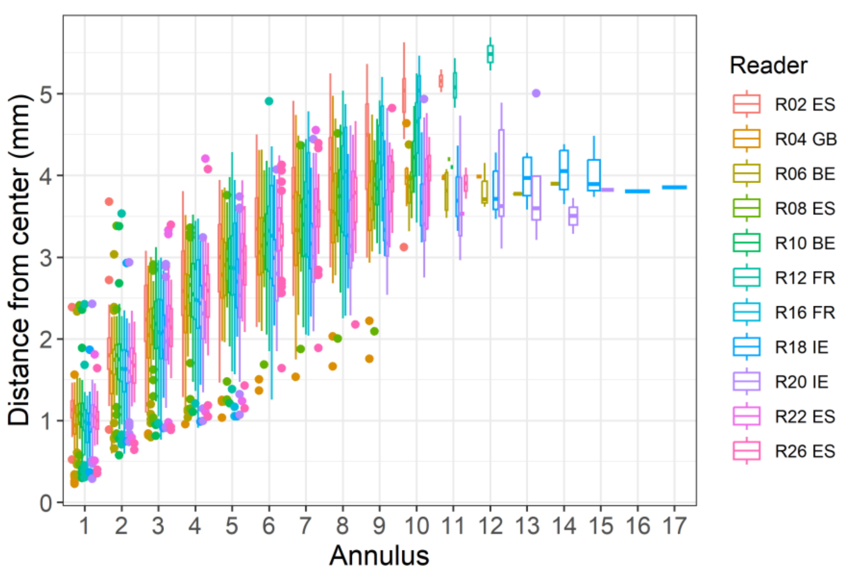

Figure 2: Plot of average distance from the centre to the winter rings for advanced readers. The boxes represent the median, upper and lower box boundaries of the interquartile range, whiskers extend no further than 1.5 * IQR (where IQR is the inter-quartile range) from the box boundary. Data beyond the end of the whiskers are represent outlers and are plotted individually.

3.9 Age error matrix AEM

Table 9: General Age error matrix (AEM). The modal age is in rows and the age classifications by the advanced readers in columns.

modal_age	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
1	0.50	0.21	0.17	0.08	0.04	-	-	-	-	-	-	-	-	-	1.00
2	0.11	0.68	0.18	0.03	-	-	-	-	-	-	-	-	-	-	1.00
3	-	0.17	0.69	0.12	0.02	-	-	-	-	-	-	-	-	-	1.00
4	0.01	0.04	0.10	0.73	0.08	0.04	-	-	-	-	-	-	-	-	1.00
5	-	-	-	0.09	0.79	0.12	-	-	-	-	-	-	-	-	1.00
6	-	-	-	-	0.11	0.69	0.16	0.04	-	-	-	-	-	-	1.00
7	-	-	-	-	0.02	0.17	0.62	0.17	0.02	-	-	-	-	-	1.00
8	-	-	-	-	-	0.03	0.13	0.64	0.18	0.03	-	-	-	-	1.01
9	-	-	-	0.02	-	0.03	0.08	0.22	0.46	0.15	0.02	-	0.02	-	1.00
10	-	-	-	-	-	-	0.02	0.13	0.22	0.48	0.07	0.06	-	0.02	1.00
11	-	-	-	-	-	-	-	0.17	0.33	-	0.33	0.17	-	-	1.00

4 Conclusion

The SmartDots Event 3985 intercalibration exercise evaluated the consistency of megrim (*Lepidorhombus whiffiagonis*) age determination among advanced readers contributing to the assessment of ICES stock 27.7b–k and 8.a, 8.b and 8.d. Analysis of advanced readers demonstrated good precision, with percentage agreement of 64%, CV of 17% and APE of 11%.

Despite the improved performance of advanced readers, complete agreement was not achieved. Inter-reader bias tests identified statistically significant differences between several reader pairs, indicating that some systematic differences in interpretation remain. Although overall bias among advanced readers was generally limited, individual readers showed tendencies to consistently over- or underestimate ages in specific age classes. The main sources of disagreement were associated with interpretation of the first annulus, identification of split or accessory rings, edge characteristics and the compressed growth increments observed in older fish.

Disagreements were primarily concentrated between adjacent age classes, demonstrating that differences generally represented minor variation in the placement or recognition of annual increments rather than major misclassification. This pattern indicates that advanced readers broadly interpret otolith structures similarly but may apply different criteria when assigning specific age boundaries. The youngest age classes, particularly age-1 fish, remained challenging due to uncertainty in identifying the first annulus, while older fish (ages 9–11) showed reduced precision because of narrow increments, overlapping structures and difficulties in interpreting marginal growth.

Agreement varied markedly among age classes. The highest precision was observed for ages 2 to 6, where agreement ranged from 68% to 79%. CV values varied widely in these age classes being age 2 the worst (29%) which improve progressively until age 5 (9%). In contrast, the youngest age class (age 1) showed the poorest performance, with PA of only 50% and CV exceeding 60% among advanced readers. The overall age error matrix indicates that age determination of megrim was generally consistent for intermediate age classes, while greater uncertainty occurred for the youngest and oldest fish. The results indicate a marked tendency to overage age-1 fish, suggesting persistent uncertainty in the interpretation of the first annulus. This ageing bias may cause recruits and juveniles (ages 0 and 1) to be underestimated and individuals to be erroneously assigned to older age classes (ages 2 and 3), leading to inflated abundance estimates for these cohorts that are not reflected in corresponding biomass trends.

Growth analyses showed similar patterns among advanced readers, although some variation in the placement of winter-ring annotations was observed. Detailed examination of problematic otoliths during the WKARMEG workshop will therefore be important for developing common interpretation guidelines and reducing remaining sources of variability.

Overall, the SmartDots Event 3985 exercise successfully identified the main areas requiring harmonization among advanced readers before the WKARMEG workshop. Future discussions should focus on establishing consistent criteria for the identification of the first annulus, distinguishing true annual rings from split or accessory rings, interpreting edge type and marginal increment formation, and improving the assignment of the oldest age classes. Continued calibration exercises, targeted review of difficult otoliths and expert consensus will further improve the reliability and comparability of megrim age data used in ICES stock assessments.