SmartDots Summary Report for the 2024 Irish Sea Sole (sol.27.7a) age reading exchange (ID 1842)

**Coordination and analysis by Karen Bekaert, ILVO**

**January 2024**

Afbeelding met verven

Automatisch gegenereerde beschrijving

# Executive summary

This age reading exchange was held in preparation for the WKBFLATFISH1 Benchmark to check the level of agreement and bias between age readers providing age data for stock assessment of sole.27.7a. 160 images of sectioned otoliths from samples collected across the four quarters of 2021, 2022 and 2023 were taken at ILVO (Belgium) and the Marine Institute (Ireland) and uploaded to SmartDots. Age distribution went from modal age 1 to 12. 6 age readers participated in the event, from which 4 were advanced (delivering data for assessments). Readers were provided with instructions to annotate all images, provide an age estimate and a quality score for their age estimation. Readers were free to draw the reading line which they found the most appropriate for age reading of each sample. Therefore, no otolith growth analysis is included in the report. The reporting module in SmartDots was used to run a standardized analysis of age reader comparison and extract a template for a full report and summary report. Results were provided to WKBFLATFISH1 2024.

To my knowledge, no previous age reading exchanges of sole 7a otoliths took place. Before the age reading exchange, the sole 7a stock assessor drew the attention towards the fact that a higher mean weight at age was obtained in Irish sole data than in the Belgian and English dataset. Therefore, extra attention was given to the results of the Irish readers in the discussion as you would expect an age reading shift towards lower ages (underestimation of age) from the advanced Irish reader to explain this result.

Based on all readers included in the age reading exchange, the percentage agreement (PA) was 85% with a coefficient of variation (CV) of 10%. This is overall a good age reading result. 88 out of 160 samples reached a percentage agreement of 100%. Based on the results of the 4 advanced readers only from Belgium, Ireland and UK the results were even improved as can be expected with a percentage agreement (PA) of 89% and a coefficient of variation (CV) of 9%. These results seem very satisfactory from an analytical point of view. When focusing on the Irish advanced age readings, there is more deviation from modal age than the age readings from the other participating countries. For the advanced Irish reader, there seems to be an overestimation of age especially in age classes 7 and 8, and a small underestimation in the older age classes 10 and 12, while there was over- and underestimation in age class 11. This might partially explain the higher weight at age, but only for the older age classes.

The main age reading issues seem to be related to the counting of false rings, interpretation of the edge and the identification of the first winter ring.

From this exchange, we can conclude that the overall results for the age readings of sole 7a are very good. There could be some improvement in agreement with the Irish readers, but the results from this exchange can only partially explain the higher weight at age from the Irish samples as no general underestimation of age could be detected. However, setting up a cooperation between age reading labs could help improve the quality of the age readings in sole 7a.

# Overview of samples and advanced readers

**Table 2.1:** Overview of samples used for the exchange event1842. The modal age range for all samples is 1-12.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Year** | **ICES area** | **Strata** | **Quarter** | **Number of samples** | **Modal age range** | **Length range** |
| 2021 | 27.7.a | Strata\_27.7.a | 2 | 40 | 2-12 | 195-495 mm |
| 2021 | 27.7.a | Strata\_27.7.a | 3 | 14 | 1-11 | 205-480 mm |
| 2021 | 27.7.a | Strata\_27.7.a | 4 | 6 | 6-12 | 395-445 mm |
| 2022 | 27.7.a | Strata\_27.7.a | 1 | 46 | 2-11 | 190-430 mm |
| 2022 | 27.7.a | Strata\_27.7.a | 2 | 10 | 4-11 | 300-430 mm |
| 2022 | 27.7.a | Strata\_27.7.a | 3 | 16 | 1-12 | 205-480 mm |
| 2022 | 27.7.a | Strata\_27.7.a | 4 | 18 | 1-10 | 180-425 mm |
| 2023 | 27.7.a | Strata\_27.7.a | 1 | 5 | 3-9 | 260-370 mm |
| 2023 | 27.7.a | Strata\_27.7.a | 2 | 5 | 5-11 | 320-430 mm |

**Table 2.2:** Overview of number of readers.

|  |  |  |
| --- | --- | --- |
| **Expertise** | **strata** | **N\_readers** |
| Advanced | Strata\_27.7.a | 4 |

# Results overview

## Multimodal cases

**Table 3.1:** 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** |
| 160 | 4 % | 1 % | 0 % | 0 % |

## 

## Age readings

**Table 3.2:** Age reading table shows the number of readings by modal age.

|  |  |
| --- | --- |
| **Modal age** | **total** |
| 1 | **32** |
| 2 | **32** |
| 3 | **72** |
| 4 | **108** |
| 5 | **112** |
| 6 | **64** |
| 7 | **44** |
| 8 | **48** |
| 9 | **28** |
| 10 | **24** |
| 11 | **52** |
| 12 | **24** |
| **Total** | **640** |

## CV table

**Table 3.3:** Coefficient of Variation (CV) table presents the CV per modal age for all advanced readers combined.

|  |  |
| --- | --- |
| **Modal age** | **all** |
| 1 | **30 %** |
| 2 | **9 %** |
| 3 | **9 %** |
| 4 | **9 %** |
| 5 | **7 %** |
| 6 | **6 %** |
| 7 | **7 %** |
| 8 | **10 %** |
| 9 | **3 %** |
| 10 | **4 %** |
| 11 | **5 %** |
| 12 | **4 %** |
| **Weighted Mean** | **8 %** |

## PA table

**Table 3.4:** Percentage agreement (PA) table represents the PA per modal age for all advanced readers combined.

|  |  |
| --- | --- |
| **Modal age** | **total** |
| 1 | **91 %** |
| 2 | **97 %** |
| 3 | **92 %** |
| 4 | **89 %** |
| 5 | **93 %** |
| 6 | **88 %** |
| 7 | **91 %** |
| 8 | **79 %** |
| 9 | **89 %** |
| 10 | **88 %** |
| 11 | **85 %** |
| 12 | **79 %** |
| **Weighted Mean** | **89 %** |

## APE table

**Table 3.5:** Average Percentage Error (APE) table represents the APE per modal age for all advanced readers combined.

|  |  |
| --- | --- |
| **Modal age** | **all** |
| 1 | **12 %** |
| 2 | **3 %** |
| 3 | **5 %** |
| 4 | **5 %** |
| 5 | **3 %** |
| 6 | **2 %** |
| 7 | **3 %** |
| 8 | **6 %** |
| 9 | **2 %** |
| 10 | **2 %** |
| 11 | **2 %** |
| 12 | **2 %** |
| **Weighted Mean** | **4 %** |

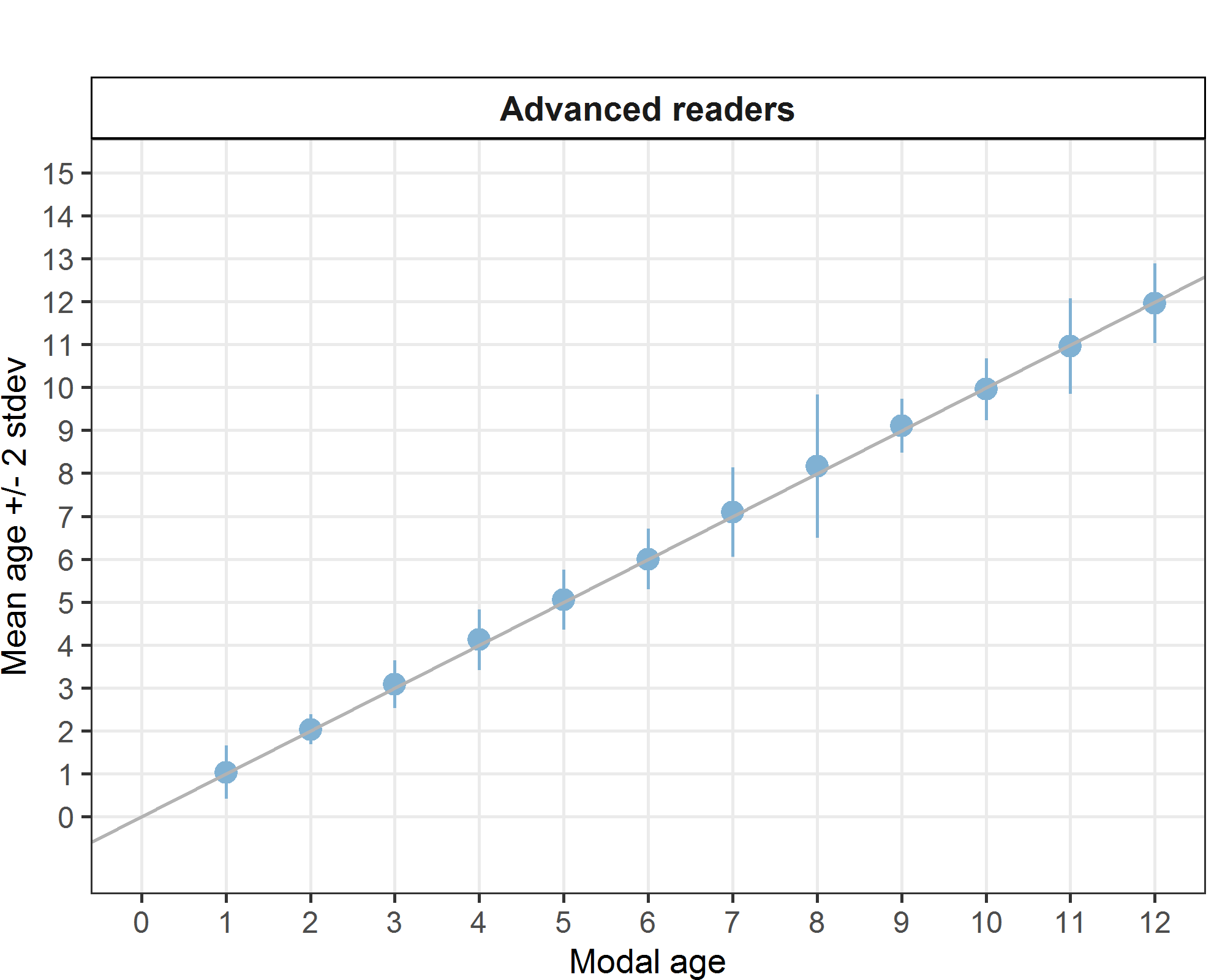
## Relative bias table

**Table 3.6:** The relative bias (as the difference between the mean and modal age) per modal age for all advanced readers combined.

|  |  |
| --- | --- |
| **Modal age** | **all** |
| 1 | **0.03** |
| 2 | **0.03** |
| 3 | **0.08** |
| 4 | **0.12** |
| 5 | **0.05** |
| 6 | **0.00** |
| 7 | **0.09** |
| 8 | **0.17** |
| 9 | **0.11** |
| 10 | **-0.04** |
| 11 | **-0.04** |
| 12 | **-0.04** |
| **Weighted Mean** | **0.06** |

## 

## Bias plot



**Figure 3.1:** Age bias plot for advanced readers.

## Age error matrix AEM for ICES area 27.7.a

**Table 3.7:** General Age error matrix (AEM). The modal age is in rows and the age classifications by the advanced readers in columns.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **modal\_age** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **Total** |
| **1** | 0.03 | 0.91 | 0.06 | - | - | - | - | - | - | - | - | - | - | - | - | 1.00 |
| **2** | - | - | 0.97 | 0.03 | - | - | - | - | - | - | - | - | - | - | - | 1.00 |
| **3** | - | - | - | 0.92 | 0.08 | - | - | - | - | - | - | - | - | - | - | 1.00 |
| **4** | - | - | - | - | 0.89 | 0.10 | 0.01 | - | - | - | - | - | - | - | - | 1.00 |
| **5** | - | - | - | - | 0.02 | 0.93 | 0.04 | 0.02 | - | - | - | - | - | - | - | 1.01 |
| **6** | - | - | - | - | - | 0.06 | 0.88 | 0.06 | - | - | - | - | - | - | - | 1.00 |
| **7** | - | - | - | - | - | - | 0.02 | 0.91 | 0.05 | - | 0.02 | - | - | - | - | 1.00 |
| **8** | - | - | - | - | - | - | 0.02 | 0.04 | 0.79 | 0.08 | 0.04 | - | 0.02 | - | - | 0.99 |
| **9** | - | - | - | - | - | - | - | - | - | 0.89 | 0.11 | - | - | - | - | 1.00 |
| **10** | - | - | - | - | - | - | - | - | - | 0.08 | 0.88 | 0.04 | - | - | - | 1.00 |
| **11** | - | - | - | - | - | - | - | - | - | - | 0.12 | 0.85 | 0.02 | - | 0.02 | 1.01 |
| **12** | - | - | - | - | - | - | - | - | - | - | - | 0.12 | 0.79 | 0.08 | - | 0.99 |

# Conclusion

Based on all readers included in the age reading exchange, the percentage agreement (PA) was 85% with a coefficient of variation (CV) of 10% and an APE of 5 %. Based on the four advanced readers providing age data for assessment from Belgium, Ireland and UK the results were even improved as can be expected with a percentage agreement (PA) of 89%, a coefficient of variation (CV) of 9% and an average percentage error of 4% which is considered very good for an age reading exchange. In general, the deviation from the modal age by an individual reader is rarely more than one year.

When focusing on the Irish age readings, it is true that there is more deviation from modal age than the readings from the other participating countries. For the advanced Irish reader, there seems to be an age reading shift towards older ages in the younger age classes (overestimation of ages, especially in age 7 and 8), but towards younger ages in the older age classes (underestimation of age). This is not what we expected: a general underestimation of age was expected as a higher mean weight at age was observed. In the older ages, there was some underestimation, but not so much. So a higher weight at age could only be explained for the older age classes.

When analysing the annotations more in detail, the main age reading issues seems to be the counting of false rings, different edge interpretation, the identification of the first winter ring, or the missing of rings when cliffed edge in older fish.

Considering the good overall results of the exchange, there doesn’t seem to be a major problem with the age readings of the sole 7a stock. However, we see some deviation in different readers, where percentage agreement is lower and bias is higher than for the other readers. Therefore, in order to try to improve the agreement on an international level, a follow-up online meeting will be held in spring 2024. Afterwards, a similar exercise as this exchange can be organised in the coming year to see if results have improved.