WORKING GROUP ON BIOLOGICAL PARAMETERS (WGBIOP)

VOLUME 7 | ISSUE 117

ICES SCIENTIFIC REPORTS

RAPPORTS
SCIENTIFIQUES DU CIEM

ICES INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA
CIEM COUNSEIL INTERNATIONAL POUR L'EXPLORATION DE LA MER

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

H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

ISSN number: 2618-1371

ICES Scientific Reports

Volume 7 | Issue 117

WORKING GROUP ON BIOLOGICAL PARAMETERS (WGBIOP)

Recommended format for purpose of citation:

ICES. 2025. Working Group on Biological Parameters (WGBIOP).
ICES Scientific Reports. 7:117. 68 pp. https://doi.org/10.17895/ices.pub.30940826

Editors

Karen Bekaert • Konstantina Ofridopoulou • Valerio Visconti

Authors

Nicolas Andrialovanirina • Gualtiero Basilone • Karen Bekaert • Cristina Bultó • Ângela Canha
Pierluigi Carbonara • Julie Coad Davies • Patricia Cortegoso • Renato Cruz • Côme Denechaud
Michelle Douglas • Laurent Dubroca • Romain Elleboode • Inês Farias • Rosalia Ferrerri
Guðrún Finnbogadóttir • Dolores Garabana • Mandy Gault • Patrícia Gonçalves • Yvette Heimbrand
Annelie Hilvarsson • Karin Hüssy • Sarah Johnstone • Léa Joly • Vasiliki Kousteni
Maria Krüger-Johnsen • Uwe Krumme • Kirsteen MacKenzie • Kélig Mahé • David Maxwell
Larissa Modica • Gráinne Ní Chonchúir • Cristina Nunes • Konstantina Ofridopoulou • Jari Raitaniemi
Joana Ribeiro • Vesala Sami • Elor Sepp • Joanne Smith • Szymon Smoliński • Lukas Snoeck
Remigiusz Szymański • Pedro Torres • Cindy van Damme • Ralf van Hal • Andrés Villaverde
Valerio Visconti • Francesca Vitale

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

Contents

i Executive summary

The Working Group on Biological Parameters (WGBIOP) focuses on reviewing the status, issues, developments, and quality assurance of biological parameters used in stock assessments. The WGBIOP made substantial progress in 2025 across all terms of reference.

ToR a addressed age-related stock issues for upcoming benchmarks, updated the comprehensive table of exchanges and workshops, and pursued quality assurance through the planning and follow-up of age calibration events. These ongoing efforts ensure that benchmark and assessment needs are aligned with group capacity, with results to be presented at the next annual meeting.

ToR b saw notable advances in the adoption of the SMSF maturity scale and histological approaches for maturity validation. The group clarified maturity staging using substages, updated conversion tables and expanded its network of experts. This facilitates a more robust and standardized approach to maturity data, while anticipating further integration of histological methods in upcoming years.

Under **ToR c**, the group monitored emerging biological methods, gave an update on the WKETAC workshop, updated guidelines on standardization, and maintained strong communication with related working groups to encourage innovation and readiness for new techniques.

Progress under **ToR d** includes facilitating the use of age error data in stock assessments, the development of two new SmartDots report types: a finalized R-scripted master table report and a simplified version for age readers with defined content. Additionally, a test case for optimizing sample selection in age calibration was identified and will proceed in collaboration with WGBIOPTIM and the Smart4SAM project.

ToR e focused on reviewing and updating linkages between environmental and biological parameters and documenting case studies, with tasks now concluded. The group also recommended the condition factor for use in contexts where age or maturity data are limited. Importantly, the review of data quality and accessibility under ToR e concluded that ongoing standardization and monitoring of biological parameter quality are best addressed under ToR a in the next year.

In **ToR f,** progress was made on training reference collections, with agreement to focus on new SmartDots-based collections (one per species). A framework outlining sample properties and precision levels was established. Input was given for SmartDots development to be carried out under the Smart4SAM project, and the next step is to implement training and quality control reference collection in practice.

Keywords: biological parameters, age, maturity, SmartDots

ii Expert group information

Expert group name	Working Group on Biological Parameters (WGBIOP)
Expert group cycle	Multiannual fixed term
Year cycle started	2023
Reporting year in cycle	2/3
Chairs	Karen Bekaert, Belgium
	Konstantina Ofridopoulou, Greece
	Valerio Visconti, UK
Meeting venues and dates	Term 2/3: 6–10 October 2025, Boulogne-sur-Mer, France and online (48 participants)

1 Introduction

Working Group on Biological Parameters 2025 report

The Working Group on Biological Parameters (WGBIOP) 2025 report presents an overview of the group's activities, achievements, and priorities during the second year of the 2024–2026 multi-annual cycle. It describes how WGBIOP continues to review, improve, and ensure the quality of biological parameters that are essential to robust stock assessment and fisheries management, while strengthening the links between data collection, calibration work, and end-user needs. The group has maintained its focus on methodological development, standardization, and the operational use of biological data and tools across the ICES area and collaborating regions.

In 2025, progress was achieved on several fronts. Work continued to improve coordination of age and maturity calibration events through the historic and new master tables and to prepare a more integrated, database-oriented framework that aligns with SmartDots and other ICES systems. WGBIOP also followed benchmark-related issues, updated the overview of exchanges and workshops, and refined communication channels with stock assessors and other expert groups to ensure that quality assurance activities respond to concrete assessment needs.

The group further advanced standardization of maturity information, promoting the use of the SMSF scale, updating conversion tools from national scales, and expanding collaborative networks for histology-based validation. At the same time, WGBIOP continued to explore and document emerging technologies, particularly AI and image-based approaches for age, maturity, and stock identification, developing draft guidance that emphasizes transparent methods, reproducible workflows, and appropriate performance metrics. These efforts build on and contribute to wider ICES initiatives on trustworthy AI and innovative biological methods.

Important steps were taken to strengthen the use of age error information in stock assessments through improved liaison with benchmark and assessment groups and the development of new SmartDots report types tailored both to analysts and age readers. Work on additional biological and life history parameters, such as condition and other traits relevant to ecosystem-based advice, led to recommendations on how best to embed quality assurance aspects within WGBIOP's core remit in future years. The group also consolidated concepts and practical arrangements for training and quality control reference collections, in close connection with the Smart4SAM project and ongoing SmartDots developments.

In 2025, WGBIOP successfully maintained a hybrid meeting model, allowing both in-person and online participation, overcoming the associated logistical challenges, and fostering a more inclusive environment for members across different regions. The next annual meeting is planned to build on the foundations established in 2024–2025, with particular emphasis on implementing the evolving data management framework, finalizing guidance documents, and further integrating reference and quality control collections into routine practice.

2 Progress report on ToRs and work plan

2.1 ToR a – Quality assurance

2.1.1 Progress during WGBIOP 2025

This ToR is a generic ToR for the group and forms part of the WGBIOP remit. This year, the subgroup working on this ToR addressed the following points:

- The identification of age/maturity-related issues of stocks up for benchmark in 2025 and onwards.
- Both the 'Historic Master Table of EX and WK' and the 'Results of age reading exchanges and workshops' tables were updated during the 2025 WGBIOP meeting. A discussion on the future of these tables was had, and proposals were made which are detailed in section 2.1.4.
- The subgroup reported results from workshops and exchanges which took place in 2025, summaries of which are available in **Error! Reference source not found.**.
- Drafted resolutions for workshops and exchanges endorsed by WGBIOP (to be approved by ICES) for 2026 and beyond, which can be found in **Error! Reference source not found.**.
- The identification of priority species and stocks in need of validation studies based on the criteria and approach approved in 2022 (see section 2.1.6).
- The status and future input of WGBIOP into the Data Quality Assurance Repository were finalized.
- A full list of exchanges has been proposed this year for 2026 and beyond, with associated coordinators (see section 2.1.3).
- WGBIOP will track the progress of proposed exchanges and workshops, facilitating the appointment of chairs, dates, and locations for workshops to convene. Results will be presented to the WGBIOP meeting in 2026 for consideration.

2.1.2 Task 1 – Coordinate communication between WGBIOP members and the corresponding stock assessment coordinators/groups of interest based on the current rolling issues list and/or other issues that have emerged, and recommendations

The issue list of the stocks proposed for benchmarks in 2026 and beyond was extracted from the Stock Identification Database (SID). This list was checked for any age or maturity-related issues. Several issues were highlighted, including the following:

- **fle.27.3a4 and fle.27.2223:** Flounder (*Platichthys flesus*) in Subarea 4 and Division 3. a (North Sea, Skagerrak and Kattegat)

Issue: Improve precision of the age reading based on age-validated material (exchange of otolith images once validated).

WGBIOP response: It is unclear if there have been any validation studies on flounder; however, the most recent small-scale age calibration exchange on flounder took place in 2021 and focused on the Baltic SD 25 area. Results were not published; however, the overall agreement among readers was poor. Neither fle.27.3a4 nor fle.27.2223 appears to be subject to category 1 age-based assessments, which indicates that a reading exchange in the absence of newly validated otoliths is unnecessary at this time.

- **whg.27.47d**: Whiting (*Merlangius merlangus*) in Subarea 4 and in Division 7.d (North Sea, eastern English Channel)

Issue: Check on the use of different maturity scales: NS IBTS Q1 Survey.

WGBIOP response: WGBIOP strongly encourages all to assess maturity directly using the SMSF scale. Some countries still prefer to use their national maturity scales and translate them into SMSF before uploading. WGBIOP has created an overview of conversion tables from national scales to the SMSF. A leaflet, to be published by ICES, on the SMSF maturity scale has been prepared by WGBIOP. This gives guidance to be followed on the use of the SMSF scale, also containing a link to conversion tables from national scales to SMSF[1]. These conversion tables have been discussed with national coordinators, quality checked by WGBIOP, and are the ones to be used for conversion before uploading data.

- **sal.27.22-31:** Salmon (*Salmo salar*) in subdivisions 22-31 (Baltic Sea, excluding the Gulf of Finland)

Issue: Inclusion in the full life-history model will require updated information regarding, e.g. smolt age distributions, maturation rates.

WGBIOP response: The most recent age calibration event on this stock took place in 2021 (SmartDots event ID 357), focusing on age reading of salmon scales. Results for expert readers were very good: Percentage Agreement (PA) 82, Coefficient of Variation (CV) of 12 and an Average Percentage Error (APE) of 5. These results indicate that an age calibration event at this time is unnecessary.

- **ele.2737.nea:** European eel (*Anguilla anguilla*) throughout its natural range

Issue: Issues highlighted focus on the integration and evaluation of the impact of recruitment data and time-series. No specific issues were highlighted with age or maturity.

WGBIOP response: Despite no specific issues being raised in relation to age or maturity, in the rolling issues list from SID, an age reading exchange of known age Eel otoliths has been proposed by WGBIOP for 2026.

2.1.2.1 Suggested update for Stock Identification Database (SID)

A suggestion for the development of the ICES Stock Identification Database was submitted to the ICES representative participating in WGBIOP. It would be useful if the SID could be adapted to include the categorization of any issues identified, which would allow users to filter for specific issues such as 'Age', 'Maturity', 'Fecundity', etc.

2.1.2.2 Complete review of National Coordinators (NCs) for age and maturity in SmartDots

A full review of the list of National Coordinators for both age and maturity was conducted during the 2025 WGBIOP Meeting. The objective of this review was twofold. First, to confirm the list of National Coordinators was still relevant, and to correct and update where changes in personnel/responsibilities had occurred. Second, to attempt to refine the list of National Coordinators to just two per country per parameter (two for age and two for maturity), in an effort to streamline international communications around calibration events.

All National Coordinators listed in SmartDots were contacted to confirm whether the current list was still accurate and to ask countries with multiple National Coordinators to consider nominating just two for their country. The role of the National Coordinators is to act as a central hub for

[1] https://github.com/orgs/ices-eg/projects/38/views/1

all communication on calibration events/queries. They then coordinate the dissemination of this information to relevant parties within their respective institutes/regions.

While replies were received from many countries, not everyone responded. Some countries were unable to narrow down their list to just two National Coordinators due to logistical reasons, but in some cases, they were able to reduce the number of people listed.

An updated list of National Coordinators for age and maturity was circulated to WGSMART and ICES to facilitate updates in SmartDots. The role of national coordinators was discussed, and a document describing the tasks is now available on GitHub[2]:

2.1.3 Task 2 – Complete annual overview of planned studies, exchanges, and workshops (i.e. update of master tables) and planning of new calibration events

2025 completed exchange and workshop reports

A total of eight age calibration exchanges and four workshops were finalized and reported in 2025, and summaries are provided in Annex 3.

Ongoing and postponed calibration events for 2025–2026

Several calibration events were planned to take place in 2025, but were either not started or are currently ongoing. These are:

- 2025 Whiting (*Merlangius merlangus*) WHG.27.7a Otolith Exchange (SmartDots Event ID: 275) exchange. Coordinator: Gráinne Ní Chonchúir. Open 16th September - 31st October 2025.
- 2025 Baltic plaice (*Pleuronectes platessa*) PLE 24 – 32 (SmartDots Event ID 2954). Coordinator: Julie Olivia Davies. Open 10th September - 30th November 2025.
- 4th European Anchovy (*Engraulis encrasicolus*) Age Reading Workshop (WKARA4). Coordinator Naroa Aldanondo. Running from the 3rd to the 7th of November 2025.
- 2025 Otoliths exchange of warm-water species in the Caribbean Sea and Indian Ocean was scheduled for 2025 but will now take place in 2026, date to be decided. Coordinators: Kélig Mahé and Valerio Visconti.

Additional Calibration Events have been scheduled for 2026.

A mix of exchanges via SmartDots and in-person workshops is proposed for 2026 and is listed below.

- European Eel (*Anguilla Anguilla*) Exchange (ele.2737.nea). Coordinator: Yvette Heimbrand. Dates to be decided.
- WKARLEM Age reading of Lemon sole (*Microstomus kitt*) IV and VIId. Coordinators: Joanne Smith and Karen Bekaert from the 30th March 2026 to 2nd April 2026.
- Workshop on Age Reading of Megrim meg.27.3a47d: *(Lepidorhombus whiffiagonis*) in Celtic Seas and northern Bay of Biscay stock (Div. 7.b–k and 8.a,b,d). Coordinators: Larissa Modica and Karen Bekaert. There will be a Small-Scale exchange between February and May 2026, with a WK in Boulogne from 22nd of June to 25th of June.
- Workshop on Age reading of Greenland halibut (*Reinhardtius hippoglossoides*; WKARGH3). Coordinator: Kristin Windsland. Dates to be decided in 2026.

[2] https://github.com/ices-eg/wg_WGBIOP/tree/master/Documents

- Workshop on Age reading 2 on Red mullet *Mullus barbatus* and striped, red mullet *Mullus surmuletus* otolith exchange (Event ID:1914 and1919). Coordinators Pierluigi Carbonara/Andrea Massaro (Italy)
- WKARHOM-5 *Trachurus picturatus*, *T. trachurus*, *T. mediterraneus*. Coordinators Andrea Massaro /Justine Diaz/ Alba Jurado-Ruzafa TBD 2026.
- Workshop on Age Reading of European Sardine (*Sardina pilchardus*). NE Atlantic and Mediterranean. [WKARAS 3] Coordinators: Cristina Bultó (IEO, Spain) and Andreia V. Silva (IPMA, Portugal).
- Workshop on Age reading of North Sea Sprat (*Sprattus sprattus*; spr.27.3a4), chaired by Julie Coad Davies, Denmark and Christine Djønne, Norway will be established and will meet at IMR, Bergen, Norway, June 1-5, 2026.
- Workshop on Age reading of whiting (*Merlangius merlangus*; whg.27.47d). Chairs: Ralf Vanhal and Côme Denechaud, Boulogne, dates TBD
- Workshop on Age reading of Irish Sea 27.7a whiting (*Merlangius merlangus*; WKISWHG) chaired by (TBC, TBC) AFBI, Belfast, Northern Ireland (Steven Beggs), Dates TBC.

The Master Table – an interactive table of exchanges and workshops

A GitHub repository (ICES_WGBIOP) has been created for all interactive tables produced by WGBIOP.

The interactive table of exchanges and workshops was updated during WGBIOP 2025 and is available on GitHub (Historic master table). This historic "master table" summarizes all completed age-reading and maturity workshops and exchanges and provides links to the published reports. However, since this table does not compile the results of the workshops and exchanges, a new table was created as discussed during the 2022 meeting (ICES WGBIOP 2022). This new table called 'Results of age reading exchanges and workshops' was also updated during the 2025 WGBIOP Meeting and now summaries the age-reading results from 2017-2025. For specific species, such as cod or mackerel as case studies, results are presented for all exchanges back in time. Since 2017, SmartDots (ices.dk) has been used for all exchanges and workshops in recent years. The results from these events are published and easily accessible within the Smartdots platform. For exchanges before 2017, only limited information is available and much more effort would be needed to compile all the results.

2.1.4 Task 3 – Review of the 'historic master table' and recommendations for future development

Background

The objective of the 'Historic Master Table' is that it acts as a repository of all calibration events for age and maturity from the late 70's to the present day. It makes it possible (in theory) to track the evolution of age reading and maturity exchanges and workshops and their outputs through the years and can be filtered to identify what stocks have poor agreement and CV's and/or have not had a calibration exercise in recent times.

There are approximately 215 stocks in historical master table and many of those still have incomplete or no data. The many challenges experienced using the Historic Master Table (it's sheer size, incomplete information, links to reports that do not work – despite having been updated several times over the years, incorrect metrics currently highlighted for maturity and a misalignment with current EX/WK outputs etc.) and the advent of Smartdots prompted the creation in 2017 of a new master table 'Results of age reading exchanges and workshops' which was better aligned with the output metrics from Smartdots (PA, CV and APE). Currently, both tables are updated by the ToR A subgroup during the WGBIOP annual meeting.

Neither Excel file currently holds information on egg and larval calibration events, or any outputs from other life-history parameters, and the maturity parameters need to be corrected.

Considerations/Recommendations

Short/Medium-Term Considerations and Goals

In the short to medium term, the objective is to end the duplication effort required to update both the historic master table and the results of age reading exchanges and workshops tables annually, as this is not an efficient use of time or resources. To do this, it will be necessary to create a single unified 'cleaned' table, where redundant variables are removed and which aligns more closely with the outputs of calibration events across all life-history parameters. This will be a task for the ToR A subgroup and a task that potentially could be started intersessionally, if time allows. Some considerations are highlighted below:

1. It is agreed that a 'Master Table' should be updated each year.
2. The historical master table needs to be adapted to align better with modern outputs from Smartdots, and the requirements to include new biological parameters. Preserving the historical data can help show that method development has given us better results, etc.
3. In relation to the historical data, this will take a great effort and cannot be achieved at the annual WGBIOP Meetings alone. Therefore, intersessional work will be required to fill in the missing data, and this can only happen if individuals are available to carry this work. Annual updates will continue to be a rolling task for ToR A.
4. The question on how to merge the information from both the historical master table and the age reading exchanges and workshops table remains and will require further discussion.
5. Proposed structuring approach: it seems sensible to consider separate tables by type of biological parameter/structure, for example, age, maturity, stomachs, eggs and larvae. We propose one file with multiple tabs.
6. These new biological parameter metrics need to be agreed upon so they can be included in the table for tracking in future.
7. Within the ToR A Subgroup, it could be possible to define the composition of these tables as well as the necessary variables, drawing inspiration from the work carried out by Remigiusz Szymański and Rebeka Tetere using the multi-sheet Excel file. That file could serve as a solid starting point for the group to review.

Each table could include:

- •the list of variables;
- •their data type (numeric, text, etc.);
- •the data source, such as:
    - via SmartDots;
    - via ICES / Stock Identification Database (SID), etc.

The table could be linked into SmartDots as well.

Regarding the use of the ICES Vocabulary, it is possible to download an updated version and to have this as a tab in the Master Table: ICES Reference Codes - RECO.

Long-Term Considerations and Goals

1. It will be necessary to move toward a database management approach, with clearly defined variables, with clear links to reference lists maintained by ICES.
2. Once the tables are merged, it may be useful to combine them with a Shiny app in R to allow easier filtering of the data, to plot results over time, etc.
3. Before this, the headings in the current master tables would need to be reviewed and edited to remove information that is not strictly needed.

4. It would be a good opportunity to address missing details, such as results by preparation method, if varying ones exist per stock and where no 'best method' has been agreed internationally.
5. Developing a handling method of changing or merged/split stocks by establishing a system that maintains links to the most up-to-date stock identification information (e.g. via SID), while preserving historical relationships between parent stocks and newly merged/split stocks. This will ensure that historical outputs remain traceable and are not lost over time
6. Ultimately, the goal is to create a database that will contain all the historical and up-to-date information on the outputs of the various exchanges and workshops across all relevant life-history parameters.

Technical Considerations and Interconnections.

Transitioning towards a proper data management approach will inevitably incur certain costs, particularly for:

- Server hosting and data storage.
- Setting up APIs or connections between databases (for example, between SmartDots/Stock Identification Database, etc.).
- Optimizing the creation form for exchanges or workshops within the SmartDots tool.
- Ongoing maintenance of the reference lists.
- Development of a suitable data extraction tool.

A suggestion is to follow a similar path to the Smart4SAM project (presented in ToR F). Eventually, it seems inevitable to move toward this type of architecture to ensure the sustainability and reliability of the data management system.

Dissemination of Information

Developing an effective strategy for disseminating historical and current information by transitioning from stand-alone files stored in GitHub to a proper, automated and searchable data management system. This will ensure broader visibility, accessibility, and usability of the data across the wider ICES community.

2.1.5 Task 4 – Review and update guidelines for calibration events and reports to be published in the ICES Library, including fixing DOI where applicable (e.g. SmartDots publications)

Exchange and workshop reporting template

During the 2024 WGBIOP meeting, it was agreed to create a WGBIOP reporting template for exchanges and workshops to streamline the process of updating the results of these events in the Historic Master Table and the New Table. Ideally, this template would be a report option within SmartDots, further streamlining the process of including event summaries in the Tables and also in the WGBIOP report annually. This proposal was highlighted in the plenary and with a member of WGSMART, who also participates in WGBIOP.

The new template was used by all EX and WK coordinators to report their results from 2025 calibration events and proved very useful. The template was also further discussed by Subgroup F and will be actioned by WGSMART.

How to publish SmartDots Reports in the ICES Library

During WGBIOP 2024, a meeting was held with the ICES Library to discuss the possibility of publishing exchange and workshop reports from SmartDots in the ICES Library. The ICES

Library will require reports in a standardized format, and these cannot be published by another entity. For example, currently, a number of the exchange or workshop reports appear to be published by individual Institutions. Once a clear path is identified, the ICES Library will revert with the relevant information, and this will be actioned.

Discussions continued throughout 2024 and 2025, but nothing has progressed due to issues beyond anyone's control, in relation to staff turnover and time to train up new joiners to the library team. The ICES Library Team have indicated that they will be able to focus on the issue of how to publish SmartDots reports in the ICES Library before the end of 2025. Therefore, it is hoped that a solution could be ready to present at next year's meeting in October 2026.

Guidelines for exchanges and workshops.

The guidelines for exchanges and workshops were also reviewed and updated during WGBIOP 2025, and some recommendations were added to coordinators to try to reach an agreement on a 'best method', where several methods are still in use for stocks and in particular institutes. These are a valuable resource for anyone wanting to set up a calibration event, including a handy checklist for both exchanges and workshops. The updated guidelines are available in ICES repository.

The WGBIOP EX and WK guidelines are also linked in SmartDots, but currently refer to the guidelines from 2023 and will need to be updated as a rolling task by WGSMART each year after the WGBIOP meeting.

2.1.6 Task 5 – Overview and recommendation of calibration exercises (age, maturity, etc.)

The validation of age estimation

During WGBIOP 2025, an exercise was carried out with the aim of identifying stocks in need of calibration (outside of the benchmark cycle) based on the criteria and approach approved in 2022. The exercise was based on and facilitated by using the data included, updated, and rearranged on the master table. In brief, the following protocol and filters were applied to identify the species and stock:

- Filter to select only species with age-based analytical assessment (column K).
- Based on percent agreement (PA in column P), fill all rows with PA < 70% red.
- Based on the coefficient of variation (CV, column Q), mark text red in all rows with CV > 15%.
- Based on the average percent error (APE, column R), make text bold in all rows with APE > 15%. (Note: not all species have APE calculated.)

The outputs of this exercise are presented below. While the list of species and stocks highlighted here might reflect the real need for a calibration exercise, this was not discussed during the 2025 WGBIOP plenary. However, if approved by the chairs, it is recommended to communicate with the age coordinators of the relevant countries/institutions to try to identify if there is capacity for anyone to take on the coordination of an age calibration exchange for these stocks in 2026.

Stock code	English name	Areas	Assessment WG	Age: % agreement from age readers, reading for assessment from most recent EX/WK	Age: CV% from most recent EX/WK	Age: APE% from most recent EX/WK	Latest EX/WK	Stock category
lin.27.5a	Ling	27.5	NWWG	48%	22%	13%	2020	1
bli.27.5b6712	Blue ling	27.5, 27.6, and 27.7	WGDEEP	34%	17%	13%	2020	1
tur.27.4	Turbot	27.4	WGNSSK	78.0	19% (stained, sectioned otolith)	7% (stained, sectioned otolith)	2018	1
hke.27.8c9a	Hake	27.8 and 27.9	WGBIE	30.7% blind and 49.7% interpretation	N.D.		2019	1
mon.27.78abd	White anglerfish	27.7 and 27.8	WGBIE	70	26	N.D.	2019	1.8
Tusk.27.5a14	Tusk	27.4 and 27.14	WGDEEP	44	15	11	2020	1
her.27.28	Herring	27.3	WGBFAS	No Data	No Data	No Data	2015	1
had.27.5b	Haddock	27.5	NWWG	No Data	No Data	No Data	2013 (AV)	1.7
had.27.5a	Haddock	27.5	NWWG	No Data	No Data	No Data	2013 (AV)	1.7
had.27.7a	Haddock	27.7	WGCSE	No Data	No Data	No Data	2013	1

A query to WGBIE regarding both hake and monkfish is included in section 2.1.6, which will confirm whether an ageing exchange is required for these species.

2.1.7 Task 6 – Recommendations and queries for WGSMART and WGBIE

Recommendations for WGSMART

- The WGBIOP EX and WK guidelines are linked in SmartDots but currently refer to the guidelines from 2023 and will need to be updated as a rolling task by WGSMART each year after the WGBIOP meeting.
- Establishing a naming convention for events in SmartDots is considered important. It is suggested that the naming convention should align with the Stock Identification Database (SID), which mirrors the headings in the master table updated by ToR A and which also aligns with the Rolling Issues List and how proposals for Benchmarks are named.
- An example would be had.27.7a Haddock (*Melanogrammus aeglefinus*) in Division 7.a (Irish Sea). Ecoregion: Celtic Seas.
- Following the above naming convention would allow WGBIOP to more easily track calibration events when reviewing the rolling issues list and the benchmarks, and identifying what stocks need to be prioritized for calibration events.
- There seems to be an issue with events being completed but final reports not being published - it would be good to consider how to address this.
- It would also be useful to be able to search the list of events in SmartDots for a stock or species, year or by ecoregion.
- There was some confusion this year in relation to events being labelled as training when they were in fact a WGBIOP exchange. It would be useful to review the option available for users to 'tag' an event and to remind users to select WGBIOP age exchange/workshop for international calibration events, and to ensure that internal, or smaller events, are highlighted as 'training', etc. Update the guidelines for setting up events in SmartDots, detailing why it's important to differentiate.

Queries for Assessment Working Groups

WGBIE: request an update on *Lophius* spp. and *Merluccius merluccius* in relation to age. These species are currently highlighted in the DCF for the collection of age-reading material (otoliths and illicia), and MS are collecting but not age reading them. Neither species is subject to an age-based assessment, so the question is whether MS must continue collecting age-restricted reading material. And if deemed necessary, what are the minimum requirements? Have there been any validation studies? Or other techniques investigated?

Recommendation to ICES for Stock Identification Database

A suggestion for development was also sent to ICES regarding the SID. It would be useful if the SID could be adapted to include the categorization of any issues identified, which would allow users to filter for specific issues such as 'Age', 'Maturity', 'Fecundity', etc.

2.2 ToR b – Promoting the unequivocal understanding and adoption of SMSF scale as well as advising on the histological approach as a validation method of maturity staging

2.2.1 Progress during WGBIOP 2025

This ToR is a generic ToR for the group and forms part of the WGBIOP remit. This year the subgroup working on this ToR addressed the following points:

1. Liaise with WGs to clarify the (sub)stages, ensuring the proper adoption of SMSF scale and receiving feedback on potential issues related to maturity.
2. Prodduce stock-specific tables with the identification of the main timing of gonadal development and spawning period supporting the proper use of the SMSF scale.
3. Correct the previously published conversion tables to SMSF scale.
4. Draft a working plan for the adoption of the histological approach as validation or estimation method, and liaise with RCGs to establish the working plan at the European scale according to the evaluation needs.
5. Encourage evenly distributed WGBIOP participation of experts on maturity and other biological parameters.

2.2.2 Task 1 – Liaise with WGs to clarify the (sub)stages, ensuring the proper adoption of SMSF scale and receiving feedback on potential issues related to maturity

Intersessional maturity experts met with (representatives of) assessment and survey groups. It became clear that more communication is needed with survey and commercial sampling groups to ensure proper adoption of the SMSF scale. To this end WGBIOP has prepared a leaflet explaining the use of the SMSF scale. This leaflet will get a DOI and be made available through the ICES library. The leaflet is intended specifically for those that assess macroscopic and microscopic maturity, but also includes a short explanation how to use the (sub)stages in different maturity analyses. The leaflet will be made available to all maturity coordinators as well as to chairs of survey groups, WGCATCH and RCGs.

Together with the leaflet a public project board has been created on the WGBIOP GitHub repository, the maturity forum. The aim is to provide a place where users can ask their fish and crustacean maturity and SMSF maturity scale related questions. Anyone with a GitHub account can post an issue in the column 'To evaluate'. A short manual on how to create a GitHub account and add an issue has been added as a pinned issue to the column 'Manual'. The maturity core group, consisting of the current ToR b members, will check on posted issues and arrange action to be taken if necessary.

Furthermore, the maturity scale conversion tables as well as the stock specific maturity table have been added to the GitHub repository Tor B folder as separate files. These are publicly available to view, but only the maturity core group (consisting of WGBIOP ToR B members) is able to modify these two files. Instructions on how to send updates for these files to the maturity core group are available alongside the file.

2.2.3 Task 2 – Produce stock-specific tables with the identification of the main timing of gonadal development and spawning period supporting the proper use of the SMSF scale

In 2024 this table was started and further updated intersessional and during WGBIOP 2025. This table is publicly available on the maturity forum.

2.2.4 Task 3 – Correct the previously published conversion tables to SMSF scale

A first draft of this table was created in 2024 based on the previous maturity staging workshops that have been held within ICES. It became clear that not all interpretations of the national scales were correct and information from some countries was missing. This table is updated and made available to the public through the maturity forum.

2.2.5 Task 4 – Draft a working plan for the adoption of the histological approach as validation or estimation method, and liaise with RCGs to establish the working plan at the European scale according to the evaluation needs

WGBIOP 2024 identified 9 steps for fulfilling this task and proposed a reasonable time frame.

1. Identify **all the stocks** with time variant (somehow) **maturity data used in assessment** by WG (start form the SID for ICES) list thus by Area (Baltic, NS, Meds etc) and which countries are taking maturity data for those shared stocks. All this info is included in ices reports and stock annex and RCGs for countries collecting maturity. *This step is finished and the table is made available, see paragraph 1.1.2.*
2. Identify experts to be involved in the definition of a statistical sound protocol for **sampling design/size** compared with the collection for macroscopic inspection. *An overview is now available of which institutes have this expertise available. WGBIOP will try to get one or more experts involved in the preparation of a sampling plan for example stock(s) by WGBIOP 2026.*
3. Identify **main laboratories able to process histological samples**: AZTI and IEO (Spain), WMR (The Netherlands), CNR and University of Cagliari (Italy), FRI (Greece), IMR (Norway), SLU (Sweden), DTU (Denmark), IPMA (Portugal), Cefas (UK), Ifremer (France). *Most institutes have provided information and protocols for histological processing of maturity samples.*
4. **Standardized protocol** (solutions, resin or paraffine, stainers, section thickness, etc). All the histolab share their protocols. *Intersessionally by March 2026.*
5. **Estimate the cost** per sample by each histolab and estimate the cost of personnel/sample by each macrolab. *Some, but not all institutes have yet provided the cost per sample.*
6. Investigate the status of **picture/scan analysis** with standardized protocols in reference to machine learning/AI. *An update of developments has been presented by Spain and the Netherlands. Intersessional a maturity AI day will be organized in 2026 to further discuss and plan developments.*
7. **Inform RCGs and involve ICES**. *This needed to be postponed to WGBIOP 2026.*
8. Plan an **ICES training course** that gather all the histologist plus other potential labs who wants to start processing own samples. Both processing and results interpretation based on the SMSF scale. We need to have ready cost analysis for building an histolab. *The*

planning of the workshop has been postponed to 2026 because first the above information needs to be available.

9. Proposal to RCGs. *After WGBIOP 2026.*

2.2.6 Task 5 – Encourage evenly distributed WGBIOP participation of experts on maturity and other biological parameters

The idea behind creating a network of national maturity experts is to get all the countries, collecting maturity data on shared stocks, involved in WGBIOP's work in order to ensure that the data quality is maintained and standardized protocols are followed across countries. Ultimately the main goal is to incentivize the participations of maturity experts in WGBIOP. ToR B members contacted maturity stagers intersessional with updates and requests regarding the maturity work done in WGBIOP. This has led to an increase in actively participating maturity experts in 2025. In 2024 nine WGBIOP members participated in ToR b. Unfortunately, four of these were not able to attend WGBIOP in 2025, of which one is no longer a WGBIOP member. However, in 2025 seven new maturity experts participated actively in ToR b, next to the four members that also participated in 2024 and earlier in ToR b.

In 2025 the contact list of maturity stagers' coordinators by country has been further updated. For two countries (Malta and Cyprus) no national coordinator is known, previous coordinators can no longer be contacted. In general, WGBIOP has drafted a definition of a maturity (and age) national coordinators role.

2.2.7 Guidelines for maturity staging workshops and image quality for SmartDots events

The guidelines for maturity stage workshops and exchanges have not been updated in 2025. There were no maturity workshops or exchanges conducted in 2025, and no recommendations were received to update the guidelines.

2.2.8 General maturity issues

2.2.8.1 Switching substages Ca and Cb not possible

In the SMSF the substage Ca actively spawning comes before substage Cb spawning capable. In the substage Cb fish are included that are between batches in the current spawning season, but also specimens that are just before spawning the first batch. For total spawners, fish that spawn all the eggs in one batch, the substage Cb comes before Ca and after that directly to Da. Thus, for total spawners the order of Cb before Ca can cause confusion when assessing the maturity stage. The WKMATCH (ICES 2014) chair confirmed that in the WKMATCH report these Ca and Cb substages had been switched in the tables. WGBIOP asked the ICES datacentre if it would be possible to switch these two stages to the more sensible order. However, data had already been entered in the ICES databases using substages Ca and Cb. It is therefore not possible to switch these stages without creating new vocabulary. Thus, in the SMSF scale the substage Ca remains actively spawning and Cb spawning capable.

2.2.8.2 Upload to the ICES databases

In 2024 it was suggested that all institutes would be allowed to upload maturity data to the ICES databases in the national scale and the conversion to the SMSF scale would be done in the database. The download would only be in de SMSF scale. The theory behind this is that the institute would have the most expertise in the national scale, and the conversion would always be done in the correct way in the database.

It was discussed again at WGBIOP and the conclusion was that WGBIOP does not support this. WGBIOP encourages all institutes to implement using the SMSF scale for assessing maturity stage. Various examples were shown in the past year where conversions have gone wrong leading to maturity ogives that cannot be used for the assessment. Also, if institutes continue to use their national scales it will not be possible to carry out reliable quality control and calibrations at workshops and exchanges.

2.2.8.3 Use of substages in the SMSF scale

WGBIOP showed in the 2024 report the importance of using the substages for macroscopic maturity staging. It became apparent in 2025 that some institutes are still uploading maturity data in stages B, C, and D. Countries not uploading substages, particularly Ba and Bb, have the responsibility to provide information to the respective stock coordinators/assessors for each stock, clarifying how their stage B should be interpreted. Specifically, they should indicate whether this stage is to be included in or excluded from the estimation of maturity ogives and in turn Spawning-stock biomass (SSB).

2.2.9 WGBIOP recommendations regarding maturity staging

WGBIOP received recommendations regarding maturity staging from WGBFAS, WKBMAC-NSSH, WGNSSK and WGDG. These recommendations have been answered and uploaded to the recommendations site.

2.2.10 Workplan for 2026

- Prepare standardized protocol (solutions, resin or paraffine, stainers, section thickness, etc) for histological preparations. All the histolab share their protocols. Intersessionally by March 2026.
- Invite experts to be involved in the definition of a statistical sound protocol for sampling design/size compared with the collection for macroscopic inspection.
- Estimate the cost per sample by each histology lab
- Organize an online meeting to discuss the status of picture/scan analysis with standardized protocols in reference to machine learning/AI, and how to align developments.
- Inform RCGs and involve ICES of proposing histological maturity staging workplan.
- Check for the possibility to create an identification key of the substages, using guiding questions (see point 6 in the maturity leaflet)

2.3 ToR c – Follow-up on emerging tools and methods for determination of biological parameters

Under this ToR, WGBIOP focuses on the monitoring of emerging tools and methods such as AI-driven technologies, epigenetics or high-dimensional shape analysis to determine biological parameters and automate/facilitate their study. The subgroup focuses on liaising with relevant actors within and outside ICES, curating a list of emerging methods and their potential applications, costs, benefits and limitations, then developing and providing guidelines and protocols for their standardization and quality assurance for a more widespread implementation.

2.3.1 Status after WGBIOP 2024

WGBIOP 2024 debuted the new resolution in which this ToR belongs. During the working group, the main progress was made on curating a list of contact points across the ICES ecosystem,

mapping the most relevant working groups and individuals for liaising with WGBIOP's work on emerging technologies. Additionally, ToR C carried out a horizon scanning of emerging methods of interest and compiled them into a living document. A takeaway from the group was the need for a workshop to bring together experts in fish biology, stock assessment, and computational technologies to explore cutting-edge methods for extracting biological information from calcified structures such as otoliths, scales, and spines.

2.3.2 Progress during WGBIOP 2025

One of the main deliverables associated with ToR C in 2025 was the organization, during the week preceding WGBIOP, of the workshop on emerging methods and technologies for the automated analysis of calcified structures (WKETAC) from October 2nd to 3rd.

During this year's meeting, the group then worked on:

- Writing the WKETAC report and summarizing its key takeaways.
- Liaising with the previously established contact list, consolidating on the WKETAC outcomes.
- Developing a set of recommendations for the successful implementation of emerging methods in biological parameters determination.

2.3.3 WKETAC

The detailed report of the workshop outcome is available in the ICES library, and only the abstract is reported here.

Every year, more than one million pieces are collected and analysed across the world to estimate individual age, which is central to most stock assessments, but also for other applications such as stock identification. Recent years have seen significant advances in the analysis and use of calcified structures for extracting biological information, and the development of new technologies related to image processing and analysis have shown a strong potential for automation and standardization. This workshop aimed to bring together institutes and experts across the world to provide an exhaustive review of the developments in this field, outline areas and stocks of highest interest, and develop an international framework for collaboration. More than 70 European and non-European researchers participated. The first day was devoted to presentations of projects, methods and case studies, which were discussed among all participants. It provided an opportunity to present different emerging methods such as AI-driven image analysis or spectroscopic data. In the case of deep learning approaches, the workshop was useful in reviewing and discussing the acquisition of two-dimensional and three-dimensional images, their standardization, and the algorithms to be developed. Discussions also touched upon the advances and limitations of these new technologies and approaches. The second day was mainly devoted to discussions around the optimization of the automated analysis of calcified structures for fish population monitoring, and how to develop a collaborative framework for data exchange and standardization. Among the output of those discussions was the creation of a dynamic repository of expertise listing, for different approaches, the associated institute, the target species and stocks, sample availability and current limitations. Additionally, the group began discussing pilot studies to successfully implement these automated methods, which will be further developed within an international project proposal to create and support a digital twin of physical archives in which large image repositories could be curated and exchanged to assist the development of automated methods and operational tools for monitoring commercial species.

2.3.4 Liaison with relevant working groups and actors

Building on the list established during WGBIOP 2024, as well as the attendees of WKETAC, the ToR c subgroup contacted the chairs of relevant working groups and upcoming workshops.

SIMWG

Chairs: Christoph Stransky (outgoing), ***Florian Berg (incoming)***

The Stock Identification Methods Working Group (SIMWG) reviews new methods for the definition and investigation of stock structure and provides recommendations to other ICES expert groups on how to interpret patterns of population structure.

SIMWG is of high relevance to WGBIOP's ToR C as it reviews the collection and use of many biological parameters to determine populations and stock identities. During WKETAC, a number of promising methods and developments were shown and discussed for calcified structures alone. The chairs were also informed about the work done by ToR C on quality and standardization guidelines for methods deriving biological parameters, which will focus on codifying quality metrics such as performance, transparency, uncertainty, or reproducibility, to ensure that said parameters can be duly evaluated before inclusion in assessment or other activities.

ToR C's initiative was welcomed by SIMWG, which was highly interested in the newer developments presented in WKETAC, such as multidimensional shape analyses or the calibration of spectroscopic tools like FT-NIRS for the identification of species and stocks. As SIMWG moves toward a new resolution with new chairs, this timely contact from WGBIOP will be useful to ensure cooperation and exchanges.

WGAGFA

Chairs: Ian Bradbury and Naiara Rodriguez-Ezpeleta

The Working Group on the Application of Genetics in Fisheries and Aquaculture (WGAGFA) focuses on the application of genetic and genomic analysis for management and conservation purposes. While there is a gradual increase in the routine use and integration of these genetic tools in assessment, recent developments on key stocks such as cod or herring in the North Sea show the difficulties associated with it. While new advances in genetics were not directly discussed during WKETAC, given its focus on calcified structures, WGBIOP may need to consider them as a biological parameter like age or maturity.

The chairs welcomed the contact from WGBIOP and saw it as a great opportunity for collaboration. They will organize for WGBIOP ToR C to give a presentation during their next meeting in May 2026, and will also use the opportunity of meeting in Copenhagen during WKCODSAMPLING to discuss it in more detail.

WGMLEARN

Chairs: Jose Fernandes (outgoing), Ketil Malde and Hassan Moustafid (incoming)

The Working Group on Machine Learning in Marine Science (WGMLEARN) reviews and identifies machine learning applications in marine science as well as emerging developments in the field of AI and ML that could be of interest. It aims to identify key challenges and provide guidance and resources for data sharing (methods, training sets, protocols).

WGMLEARN is of high relevance to WGBIOP ToR C considering the current interest in AI-driven, deep learning approaches to biological parameters, particularly those that relate to image analysis for automatic aging, maturity staging, and species/stock identification. WGMLEARN also has a dedicated ToR on trustworthy AI, which has been a key discussion point throughout the WGBIOP ToR C activities and especially during WKETAC. As new methods for determining

biological parameters that will form the basis of stock assessment are adopted, such considerations will be crucial.

ToR C's initiative was well received by WGMLEARN, who happened to be assembled in their own meeting. As the working group sat to develop the new ToRs of their next iteration, they were positive about harmonizing their work with WGBIOP's and potentially making a dedicated deliverable.

WKAAII

Chairs: Laura Uusitalo and Neil Holdsworth

The workshop on applying AI in ICES (WKAAII) will be taking place online 27-30 October 2025. It aims at a broad exploration and inventory of AI use cases in ICES, from automation and assistance in report writing, via Large Language Models, to scientific developments in data collection and interpretation, via image recognition and other similar deep-learning methods. While only some facets of the workshop will be relevant to WGBIOP activities, the overall focus on examining ethical and capability considerations with AI technologies such as human accountability, reliability, and transparency will have a significant overlap with ToR C's work.

It was proposed to WKAAII that a presentation was given on the outputs of WKETAC and the current work done by WGBIOP ToR C, to further liaise with relevant ICES actors and harmonize the development of guidelines on using AI-approaches for those specific use-cases relevant to the acquisition, treatment and analysis of data on biological parameters. The chairs were interested in the outcomes of the workshop and welcomed the suggestion, and the presentation was held during the second day of WKAAII as one of the "use cases" presented. Further, the general use case of otoliths and calcified structures with machine learning was submitted in detail for their horizon scanning of current and upcoming usage within ICES, and was further discussed in subgroup work.

2.3.5 Recommendations and guidelines for emerging technologies

Discussions in ToR C, during WGBIOP 2024, highlighted the need to push for an international effort towards standardizing and sharing input data (images, annotations, etc) as well as all acquisition, processing, and training procedures. It is not feasible to control all parameters given the breadth of instruments and practices across institutes, so the focus should be on goal-oriented guidelines (like the WGSMART guidelines for image quality). The idea was to provide guidelines and quality standards for how images should look so that a human expert can read them (centralized, large otolith, good contrast, and visible rings, smooth uniform background, etc.), then curate processing algorithms that can transform these raw images into standards to be used in training.

During WGBIOP 2025, such preliminary guidelines were drafted for age, stock, and maturity determination using artificial intelligence techniques and image analysis. The aim was to initiate a set of good practices tailored to these emerging methods and formulate a list of important issues to consider when developing official guidelines. It is important to emphasize that this work represents only a starting point and should not be considered a finalized guideline. This section of the report summarizes the key points raised during the discussions.

Existing guidelines

The subgroup recognizes that the existing WGSMART Quality Guidelines provide a solid foundation for standardizing image collections and associated data. These guidelines can serve as a valuable starting point for developing more specialised protocols tailored to image-based and AI-driven approaches. The idea is to build upon it by adapting its principles to the specific needs of automated image analysis, including techniques such as deep learning and computer vision.

The goal is not to replace the WGSMART framework but to complement it with targeted recommendations that address the unique challenges and opportunities of AI-based methods. In practice, the guidelines should focus on the standardization of sample preparation methods, method development (including detailed information on the model training phase), and evaluation of the process or algorithm in terms of performance, uncertainty, trustworthiness, and reproducibility.

Sample preparation methods

The subgroup noted that there are various sample preparation methods used for age reading, stock identification, and maturity analysis. These include different techniques such as sectioning, histological sampling, and staining procedures. All these aspects can significantly influence image characteristics and model outcomes. Therefore, they should be thoroughly documented in the metadata of any dataset used for AI model training to ensure reproducibility and meaningful comparisons across studies.

Image acquisition

For image acquisition, the subgroup advises following SmartDots guidelines as a strong foundation, but additional metadata that refer to the image, the preparation method, and the sampling information should also be recorded. First, in the case of aging structures such as otoliths or scales, they need to be isolated one by one (no overlapping), and any artefact on the image that is not the intended object, such as tissue, dust specks, air bubbles, or hair, must be avoided. This is essential because, although many labs already have algorithms that can identify an otolith or scale in the image, some algorithms cannot identify and separate artefacts from the actual object. Second, for each image, the properties of the camera devices used should be collected, as well as additional information, for example on i) how many otoliths are in the image; ii) if there is any broken otolith in the image; iii) the visible face of the otolith; iv) what kind of preparation was performed (whole, section, broken and burnt); v) the light source (transmitted or reflected); vi) the magnification; viii) the light intensity (standardized according to SmartDots instructions).

It is important to have raw format images (keep all the image information because it is unprocessed data straight from a digital camera). Having all the image data may be useful for models now or in future. However, it has been remarked that a few key features on the image can provide better results, and more data can add more noise; hence, we need to figure out which is the best balance solution. Most of the time, TIF format (Tag Image File Format) is provided for most camera devices. Moreover, images need to have the real DPI (density of pixels per inch) to have the real size of the object. Keeping a database with the RAW images, combined with the documented standardization scripts and steps undertaken publicly available for all users, will assure the transparency and reproducibility of the method.

Image databases

When designing a training database, it is important to implement best practices and build on previous experience from similar initiatives. Projects like FathomNet offer valuable examples of how to structure metadata, manage image diversity, and support reproducibility in AI-driven biological analysis (Katija *et al*., 2022; Jan *et al*., 2024). Leveraging experiences from the development of existing otolith image datasets would also be useful, especially those that have addressed similar challenges in metadata structuring, image variability, and reproducibility. A database of around 12,000 unstandardized two-dimensional images of plaice and striped red mullet otoliths is available in SEANOE (Andrialovanirina *et al*., 2022). These images are provided as support for a publication describing an automated pipeline for the processing and standardization of otolith images for age reading, shape analysis, and other image-based approaches. Similarly, a repository of more than 330 striped red mullet otolith three-dimensional reconstructions

(Andrialovanirina *et al*., 2024), as well as three-dimensional images from 22 common fish species in French fisheries (Andrialovanirina *et al*., 2025), are hosted at SEANOE. Built over two decades, the AFORO collection (Lombarte *et al*., 2006) is one of the largest open-access otolith collections, featuring over 9,000 two-dimensional and three-dimensional images from many species.

Image resolution

The subgroup emphasized that image resolution plays a critical role in model performance. Many existing models rely on low-quality images where otolith annual increments are barely visible. Recent findings (e.g. Miyajima-Taga *et al*., 2025) show that higher resolution improves results, but only up to a point beyond which gains become marginal due to noise and signal limitations. This supports the need to optimize image standards, balancing performance with cost and practicality. Additionally, some models are designed to handle flexible input resolutions, offering adaptability across different imaging setups (e.g. Cayetano *et al*., 2024).

Pre- and post-processing

The subgroup also recognizes that model performance is influenced not only by image quality but also by the pre- and post-processing steps applied to the images before they are used in machine learning workflows. For the sake of reproducibility and fair benchmarking, it is essential to document these steps, such as cropping, filtering, normalization, or contrast enhancement, as they can significantly affect model outcomes and comparability across studies.

Open-source programming language

The subgroup suggests adopting a common open-source programming language for scripts, since this may help standardize procedures and speed up the adoption of new methodologies by new users. Among the most promising, the subgroup identified the general-purpose programming language Julia since it appears more flexible and similar to R and Python. Julia has built-in support for calling C or Fortran language libraries and also allows interoperability with languages such as Python, C++, Rust, R, Java, and SQL. However, since actually the most used scripts are written mainly in R (25% of data scientists prefer R for statistical analysis in AI projects) or Python (70% adoption in AI and machine learning projects), they would be recommended but thinking and moving gradually to a more performer language as Julia (Julia's usage has grown over 20% in AI research papers since 2022).

Standardization methods

For all algorithms, experts could follow their own pipeline of data standardization and feature extraction or use one that is recommended and already validated by the community. In either case, all data and scripts used must be available, whenever possible, for reproducibility and transparency.

Statistical evaluation metrics

It is important to recognize that traditional age reading assessments (e.g. through exchanges of calcified structures or images) and AI model training rely on fundamentally different evaluation frameworks. In age reading, performance is typically assessed using metrics such as PA (percentage of agreement among readers), APE (Average Percent Error), CV (Coefficient of Variation), bias, and precision between readers or laboratories. In contrast, AI models, especially those used for image-based classification or regression, are evaluated using a broader set of statistical metrics. Accuracy alone is not sufficient. For classification tasks (e.g. stock discrimination), metrics such as F1-score, specificity, sensitivity, and confusion matrices are often more informative. For regression problems (e.g. age prediction), metrics like RMSE (Root Mean Square Error), MAE (Mean Absolute Error), and R^2 are commonly used.

Model performance schemes

Equally important is the way the model is trained and validated. The size and diversity of the training dataset, as well as how it is split into training, validation, and testing subsets, can significantly influence model performance and generalizability. Moreover, the choice of validation strategy, such as Leave-One-Out Cross-Validation (LOOCV) or k-fold cross-validation, plays a crucial role in the interpretation of results. It should be noted that methods like LOOCV often yield more optimistic performance estimates compared to, for example, threefold validation, which may better reflect real-world variability (Smoliński *et al*., 2020). These methodological choices must be clearly documented and considered when comparing AI-based approaches to traditional methods and also when comparing different AI-based models.

Something like evaluating for a given process or algorithm: performance, uncertainty quantifying, trustworthiness, reproducibility, etc. SWOT (strengths, weaknesses, opportunities, and threats) analysis should be considered down the road, but it is much more time-consuming and precise in its execution, as it would need in-depth cooperation between scientists, data experts, etc. Given the expected cost improvement and scalability of the emerging methods, the subgroup also highlighted a need for a rigorous economic cost analysis comparing different approaches, including the cost of human labour, energy, consumables, etc.

To ensure consistency and comparability across different models, all experts should be required to disclose the full set of model parameters and to evaluate their models using the same optimization and performance metrics defined in the statistical evaluation metrics section. Ideally, the entire analytical pipeline from data preprocessing to results should follow a standardized framework. Within this unified pipeline, only the dataset, model architecture, and model parameters would need to be modified or optimized by each expert. Such standardization would not only facilitate reproducibility but also make it easier to implement transfer learning, whether by reusing preprocessed data or transferring learned model weights across tasks. In the case of Deep Learning models, experts could further enhance transparency by providing the learned feature weights or attention maps, offering insights into which inputs most strongly influence model performance and allowing more interpretable model comparisons.

Samples included in the training and evaluation process

The subgroup discussed the training strategy dilemma: whether to use only high-quality, unambiguous images or include all available samples. While starting with clear, high-resolution images may be beneficial during early development phases, especially to teach the model basic recognition patterns or test alternative initial models, more realistic evaluations should be based on full datasets that include challenging cases and lower-quality images, reflecting typical field conditions. This aspect was highlighted by the subgroup as an important consideration when benchmarking models.

Other aspects to be evaluated

When assessing AI models for age and maturity determination, it is essential to look beyond performance metrics. While these metrics give a general sense of how well a model performs, they do not capture other critical aspects, such as transferability (the ability of a model to be applied to different datasets, species, or regions, or to be retrained efficiently for new use cases). Equally important is transparency, which refers to the interpretability of the model and the level of trust users can place in its outputs. Models that mimic expert reading patterns (e.g. focusing on specific otolith features) tend to be better received than opaque, black-box approaches such as full-image regression, even if the latter show strong numerical performance.

To support fair and meaningful comparisons between methods, the subgroup proposes considering a multidimensional evaluation framework that includes performance, transferability, transparency, and uncertainty. For example, a model might score 5/5 in performance but only 2/5 in transparency due to its complex and non-interpretable architecture. Conversely, a simpler

model might achieve only 3/5 in performance but 5/5 in transparency, making it easier to refine and integrate into expert workflows.

Benchmark data

To support meaningful comparisons and future development, it would be valuable to establish a standardized benchmark dataset for age, stock, and maturity determination (and potentially other issues). This would allow consistent evaluation across models and facilitate meta-analyses to identify key factors influencing performance, such as image resolution, model architecture, or training strategy.

Assessment-oriented evaluation

In many cases, AI models will ultimately be evaluated based on the final product, i.e. the input they provide to stock assessment models. Evaluation of algorithm performance may therefore focus on those groups most relevant to the end product, such as age/length groups used in stock assessment models. This means, for example, that in assessments where plus groups are used, distinguishing between older age classes may not be critical. However, accuracy in these older groups can still impact biological parameters such as growth curve fitting or natural mortality estimates. These nuances should be considered when evaluating and comparing models, ensuring that performance is aligned with the specific needs of the assessment framework.

Biological concepts in age reading

The subgroup discussed cases where the goal is to train the algorithm to act like a human reader by detecting, counting, and interpreting otolith rings. For age estimation, it is essential to define clear interpretation criteria that reflect how age reading varies across stocks and seasons, ensuring alignment with expert practices and assessment needs. In such cases, adding biological rules and conventions (e.g. species, stock, adopted birthdate, fish capture date) is relatively straightforward and can be incorporated as metadata during training. Alternatively, a second modelling layer could be built on top of the pattern recognition phase. Ultimately, a combination of biological factors and evaluation metrics will be key to producing reliable and interpretable outputs.

2.3.6 Workplan for 2026

Throughout 2026, many members of ToR C and attendees of WKETAC will be coming together to take the workshop outcomes and develop a joint proposal, to be presented at WGBIOP 2026 at the end of this resolution cycle.

2.4 ToR d – Review the procedures used in calibration events and facilitate the transfer of error data into stock assessments

2.4.1 Progress during WGBIOP 2025

Each of the following sections describes the progress under each task of ToR d

2.4.2 Task 1 – Liaise with relevant assessment working groups to facilitate the use of age error in stock assessment

The main actions carried out include:

- Continued communication with the relevant stock assessment working groups and benchmark groups on results of age reading calibration exercises and age data quality. Carried out the form of presentations at the assessment working groups (via the WGBIOP presentation), report outputs from SmartDots or direct communication. Table 2.4.2 gives an overview of communication.
- During the data compilation workshop for WKBHERSPR (Benchmark Workshop on selected Herring and Sprat stocks) the results of the 2025 Sprat exchange (spr.27.3a4) https://smartdots.ices.dk/ViewEvent?key=1910 were presented and discussed. Additionally, a working document was provided for the final benchmark report. The decision was made to not spend the time adapting the new model to incorporate the age error data, given the time restrictions coming up to the benchmark. Concerns were also raised over the small number of examples where the incorporation of age error data into the stock assessment procedure has been successfully implemented. However, it was agreed that this would be worked on in future and that additional age error data (from more years) would aid in model testing. These data could be provided from a new SmartDots event. Given the low level of agreement and uncertainty around the ages of the youngest fish in the exchange, a formal recommendation for an age reading workshop, via SmartDots, was made to WGBIOP by WKBHERSPR. This was answered during WGBIOP and a workshop resolution drafted for approval. The workshop is due to take place in June 2026.
- In relation to Western Baltic herring (her.27.20-24), the results from the 2024 WBSS herring exchange (ID 1853) were not presented due to time restrictions and the acknowledgement that the results will not be considered for the current benchmark. A PowerPoint presentation was made available on the SharePoint
- During the presentation given to WGBIOP from the RDBESGOV a request was made to have WGBIOP added to the list of ICES Expert Groups with access to RDBES. This will strengthen WGBIOP's capabilities to support data quality assurance for the advice process.

2025	Report Published in SmartDots	Stock up for benchmark	Presentation or Communication of results to stock assessment WG	Age error matrix or age error data provided to stock assessment WG	SmartDots summary report provided to stock assessor	Feedback from stock assessment WG	Comments
NEA Mackerel (Event ID:1888)	Y						
2025 Solea solea 27.8ab (Event ID:1903)	N						No exchange with the assessment working group to date. Latest benchmark: WKBFLATFISH 2024.
North Sea Sprat (spr.27.3a4) age reading exchange (Event ID:1910)	Y	Y	Yes, a WD and presentation for the WKBHERSPR data compilation workshop	Yes	Yes	Assessment model is being updated, age error data will be tested in the new model, post benchmark	A recommendation for an age reading workshop was made to WGBIOP from WKBHERSPR, this will take place in June 2026.
WKARBLUE4 *Micromesistius poutassou* (Event ID:1922 and 2947)	N	Yes (data preparation meeting end of 2026/benchmark 2027)	Yes	Not yet, but will be. Tests will be made to incorporate in SAM for BW SA.	No	Yes	
Gulf of Bothnia herring her.27.3031 (Event ID: 1849)	N	N	not yet	Work in Progress	Work in Progress	Not yet	Meeting planned with stock assessors and age readers in November
Lemon Sole IV andVIId (Event ID: 115)	N	Y	not yet				A lemon sole age reading workshop will be organized in March 2026, stock is up for benchmark

2025	Report Published in SmartDots	Stock up for benchmark	Presentation or Communication of results to stock assessment WG	Age error matrix or age error data provided to stock assessment WG	SmartDots summary report provided to stock assessor	Feedback from stock assessment WG	Comments
							WKBCAT1 in autumn 2026
Whiting North Sea 47d (Event ID:2941)	N	Y (2026)	Not yet	Work in Progress	Work in Progress	Not yet	Data compilation is only end of November 2026
WKARHOM-5 Exchange: *Trachurus picturatus, T. trachurus, T. mediterraneus* (Event ID:2926, 2927 and 2928)	N						
WKVALMU_2 Red mullet *Mullus barbatus* and striped red mullet *Mullus surmuletus* otolith exchange (Event ID:1914 and1919)	N						
2025_Sardine_ICES_GFCM (Event ID: 2950)	N	Y	N	N	N	N	Workshop recommended in 2026 to calibrate readers from emerging stocks and ensure adequate advanced-reader coverage — particularly in Atlantic subareas 27.9.a and 27.8.c

The annual presentation from WGBIOP to the stock assessment EG's had varying levels of success. The slides included results of relevant age and maturity exchanges and workshops, updates on the use of age error data in stock assessment and updates on the use of the SMSF scale. When the presentations were given, they were found to be informative and lead to fruitful discussions while raising the awareness of the work carried out under WGBIOP on data quality for stock assessment purposes. This will continue in 2026, with presentations also to be given to the survey working groups where any updates on the implementation of the SMSF scale will be most relevant.

Discussions on the past and present requests from WGBIOP for information to be provided from SID (e.g. stock coordinator and assessor, age plus group, which countries are providing biological data and at what spatial and temporal scale, etc.) lead to the conclusion that the ICES Data Officer for SID will provide what is needed for WGBIOP on an annual basis. This should be worked on intersessionally and a standard template agreed upon by the end of 2026.

The use of the benchmark rolling issue lists as a means of communication between WGBIOP and the stock assessment EG's does not appear to be successful. Stock assessment working groups were not informed of this in the presentation given at the start of the assessment EG meetings by the ICES Secretariat. The rolling issue lists are however being used in preparation for benchmarks but with little awareness that WGBIOP uses these lists to prioritize its cycle of exchanges and workshops. The Benchmark Stock Rolling Issue list webpage is under development in line with the ICES Benchmark webpage https://www.ices.dk/about-ICES/how-we-work/Pages/benchmarks.aspx where it is now possible to get the most up to date information on past, present and upcoming benchmarks.

2.4.3 Task 2 – Revise statistical procedures to improve the reporting of calibration events, e.g. by taking into account the age plus groups, and quality scores (AQ or QS), and investigate temporal and spatial stratification of samples at national and/or regional level

In 2019 an official recommendation was made from WGBIOP to PGDATA and RCG's "for all national laboratories to implement the AQ codes agreed upon by WGBIOP to be used for recording age quality in SmartDots and other ICES/GFCM databases". The AQ codes **https://vocab.ices.dk/?codetypeguid=e08ec685-61f6-4ccb-9e93-594047b05797** are based on those developed by WKNARC, 2011 to reflect the quality of the biological data used for stock assessment purposes. The recommendation was made so that the codes should be applied uniformly to both SmartDots annotations and on the data reported for stock assessment, thus allowing for a streamlined evaluation of data quality for stock assessment purposes. Contact was made to the RDBES data manager who provided an overview of which countries are reporting AQ codes and which codes are reported. It is apparent that the reporting of AQ codes across institutes is not standardized. In the WGBIOP 2024 report is a finalized table "Quality Status of Age Reading at Institutes", which is a compilation of quality management procedures carried out at national ageing laboratories. It gives information on how each lab is implementing AQ codes. Both sources of information clearly show that implementation is not standardized. The next step is to understand how the data submitters and stock coordinators apply the AQ codes, if they are available. This will be taken up at WGBIOP 2026. It will also be investigated in relation to maturity staging.

With respect to investigating temporal and spatial stratification of samples at national and/or regional level, the request made from WGBIOP to the RDBESGOV chairs to have WGBIOP added to the list of ICES Expert Groups with access to RDBES will aid WGBIOP in understanding and creating overviews of national and regional stratification of samples used for stock assessments. These data are already freely available from DATRAS. How these data will be used to revise statistical procedures, in light of spatial and temporal stratification of samples, and improve the setup and reporting of calibration events will fall under the EU Smart4SAM project and the 2027 Workshop on the Application of Ageing Error in Stock Assessment (WKAESA).

In relation to report outputs, the main actions carried out include:

- In cooperation with WGSMART and with input from a small number of age readers, a short and concise SmartDots report for age readers and age reader coordinators has been developed https://github.com/ices-eg/SmartDots/issues/377. It has now been sent out to all national age reader co-ordinators with feedback and input requested so that a final reviewed version is ready for the WGSMART Q4 meeting. Once finalized, it will be part of the SmartDots report zip file provided on request when the event coordinator is producing the reports. A question of whether or not this could be provided for each individual reader will depend on new functionalities being developed in the ICES SmartDots webpage. This could be similar to a report output developed under the future Reference Collection module.
- A suggested short report template from ToR a was reviewed which provides the information required for the WGBIOP Master Table. The r-script was drafted during WGBIOP, in cooperation with WGSMART https://github.com/ices-eg/SmartDots/issues/350, presented and agreed upon in plenary. Once finalized, it will be part of the SmartDots report zip file provided on request when the event coordinator is producing the reports.
- Following a suggestion from WGBIOP an overview of SmartDots reports and what output the various reports contain was compiled and presented. (Table 2.4.3)

Table 2.4.3 Overview of SmartDots report and contents

		SmartDots Full Report	SmartDots Summary Report (for stock assessors)	Master Table Report	SmartDots Age Reader Report
Overview of event (species, stock, ICES assessment EG, etc.)				x	
Methods Section		x			
Overview of samples and readers		x	x (only advanced readers)	x	x
Multimodal cases		x	x (only advanced readers)		
Summary Statistics Table		x		x	x
Number of age readings by modal age			x (only advanced readers)		
CV per modal for all readers combined			x (only advanced readers)		
PA per modal for all readers combined			x (only advanced readers)		x
APE per modal for all readers combined			x (only advanced readers)		
*CV per reader and modal age	All readers	x			
	Adv. readers	x			
*PA per reader and modal age	All readers	x			
	Adv. readers	x			
*APE per reader and modal age	All readers	x			
	Adv. readers	x			
*Relative Bias	All readers	x			

		SmartDots Full Report	SmartDots Summary Report (for stock assessors)	Master Table Report	SmartDots Age Reader Report
	Adv. readers	x			
*Bias plots	All readers	x			x
	Adv. readers	x	x (only combined for advanced readers)		
*Inter Reader Bias Test	All readers	x			
	Adv. readers	x			
*Growth Analysis	All readers	x			
	Adv. readers	x	x (only combined for advanced readers)		
AEM	Adv. readers	x	x		
* AEM by ICES Area			x		
Comparison of CV for all readers combined by strata			x		
Comparison of PA for all readers combined by strata			x		
Comparison of APE for all readers combined by strata			x		
* Option to produce by strata					

While compiling the WGBIOP Reference Collection Guidelines under ToR f and considering precision thresholds to be used when selecting samples for "consensus age" collections, concerns over the calculation of CV in the reporting module were raised. Values reported from SmartDots events are most often over the suggested 5% threshold expected from advanced readers, but it is important to note that thresholds are species and stock dependent given differences in age range and determination difficulty (Campana, 2001). This has been taken up with WGSMART who will investigate further and report back to WGBIOP.

2.4.4 Task 3 – Identify potential sources of errors in the procedures involved in generating Error Matrices (EM) or the raw data supplied to the stock assessors

- The number of samples currently used in age intercalibration exchanges is consistently determined by the coordinator, due to the absence of formal guidelines. This decision is primarily based on prior knowledge that sample distribution should encompass the relevant ICES areas and quarters, as these are used in the assessment. However, beyond this general principle, there are no specific guidelines regarding the number of samples per ICES area, quarter, or length class to be included in the exchange.
- In order to evaluate the impact that varying sample sizes may have on the estimated Age Error Matrices (AEMs) resulting from the exchanges, a sensitivity analysis will be conducted using a selected Event ID with high sample coverage (Event ID 1903). This event has been chosen as the case study due to the limited number of ICES areas associated with these stocks. The primary approach will involve performing bootstrap simulations using different subsample configurations, allowing for comparison of the resulting AEMs across each simulation setup.
- The comparison of AEMs should be based on robust statistical tests and a range of quality indicators. To support this, a recommendation has been prepared for WGBIOPTIM, seeking input from the working group on the most appropriate statistical models and metrics to include in the comparison. The goal is to determine the optimal sample size for age reading intercalibration exchanges, ensuring that the resulting AEMs are both precise and representative of the stock.
- The first step will be collaboration with WGBIOPTIM. The second will involve applying the Age Error Matrices (AEMs), derived from the optimal sample size simulations, in the assessment process, which will be carried out under WKAESA.
- This task will be integrated into Smart4SAM in future, where additional developments and guidelines will be established to determine the optimal number of samples required for estimating the Age Error Matrix used in stock assessment models.

2.4.5 Task 4 – Develop an improved sampling design for SmartDots calibration events to account for these sources of errors

Tasks 3 and 4 are strongly interlinked and will both fall under the EU project Smart4SAM - SmartDots for Monitoring, Accuracy and Reliable Training of essential biological data - Quality assurance for stock Assessment. WP4: Application of SmartDots data output in stock assessment - aims to develop methods for incorporating and testing the impact of age interpretation error data from SmartDots into stock assessment models. Best practice guidelines for sample selection for SmartDots events (to create AEM for stock assessment) will be developed under this WP and will then be incorporated in the guidelines for the organization of age reading exchanges and workshops (under WGBIOP ToR a). The project will run from November 2025 to April 2028.

2.4.6 Task 5 – Organize a second workshop on the use of Ageing and Maturity Staging Error Matrices in Stock Assessment (WKAMEMSA)

Under Smart4SAM WP5, the ICES Workshop on the Application of Ageing Error in Stock Assessment (WKAESA) will be held in November 2027. It was agreed that this workshop will focus solely on the use of ageing error matrices in stock assessment, as it is too preliminary to consider maturity staging error data in this context. It is envisioned that the workshop will use selected case studies to test the impact of including age interpretation errors on stock indices, testing data inputs and model configurations and will serve as a forum for the dissemination of results, including best practice guidelines for sample selection.

The draft resolution written at WGBIOP 2024 will be revised in 2026 based on the progress of this ToR within the current WGBIOP 3-year term.

2.4.7 Workplan for 2026

Task 1

- Continue to liaise with relevant assessment working groups to facilitate the use of age error data in stock assessment
- Compilation of previous years' WGBIOP requests for information from SID, deciding on what is necessary and relevant, and compilation of the final list to be agreed with WGBIOP members and the ICES Data Officer responsible for SID.

Task 2

- Follow-up on the standardization and implementation of AQ codes for recording biological data quality – how are these being applied at national labs, and how are they being applied by the data submitters and stock coordinators
- Continue to work with ICES WGSMART on SmartDots report content while keeping developments in line with EU Smart4SAM.

Task 3

- Collaboration with WGBIOPTIM will be established to develop and implement the most suitable statistical methods for determining the optimal number of samples to include in the age reading exchange. This will be guided by the analysis and comparison of Age Error Matrices (AEMs) resulting from the exchange using different sample sizes.

Task 4

- Liaise with ToR f in providing input to EU Smart4SAM, who will develop best-practice guidelines for sample selection for SmartDots events (to create AEM for stock assessment)

Task 5

- The resolution for the ICES Workshop on the Application of Ageing Error in Stock Assessment (WKAESA) will be drafted based on the resolution written at WGBIOP 2024 for the second workshop on the use of Ageing and Maturity Staging Error Matrices in Stock Assessment (WKAMEMSA). This will be done in cooperation with EU Smart4SAM and based on the progress of this ToR within the current 3-year term of WGBIOP.

2.5 ToR e – Potential role of additional biological and life-history parameters in stock assessments and fisheries advice and evaluation of significant changes in time

2.5.1 Introduction

ToR E focused on how additional biological and life-history parameters are integrated into stock assessments and fisheries advice. The group reviewed last year's objectives (2024) regarding the research linking environmental conditions, particularly temperature, to biological responses in demersal and pelagic fishes and invertebrates across the ICES and GFCM regions.

We consider that **relationships** between the environment and biological parameters are **well-known** within the ICES community, and that updated reviews of these relationships, except in specifically defined cases, are not necessary at this time. We again emphasize that integrating biological parameters like **body condition factor** into benchmarks and stock assessments alongside standard graphs could be **useful** and enhance ecosystem-based fisheries advice. In future, as the majority of tasks in ToR E are overly broad in scope, we suggest that **ToR E should be closed**, although the suggestions made to date within these tasks should be retained. From 2026 onwards, the quality assurance aspect of Task 4 could be integrated within ToR A of WGBIOP.

2.5.2 Task 1 – Review the links between biological and life-history parameters and (climate-induced) changes in environmental conditions

In last year's report, a first attempt was made to review this very broad topic, presenting a somewhat arbitrary overview of the studies known or easily found through literature searches by the experts present in the subgroup. Evaluating last year's work, it was concluded that, with the available time and the expertise in the subgroup and the global scope of the issue, a further attempt to populate the tables or specific search for an additional link with one or more environmental conditions would not strengthen the message, namely that **environmental conditions clearly affect the biological and life-history parameters of commercial stocks** (and their non-monitored prey species) and are therefore relevant to consider, especially under the current situation of rapidly changing environments due to climate change.

A key result of our discussions was that the requested overview within this ToR of the links between environmental data and biological/life history parameters is too broad. As this Task is currently written, the topic is considered too broad for practical use within the framework of WGBIOP. Unless significant changes in the specificity of the task are announced, we recommend that it should be closed.

Links between biological and life-history parameters and the environment could be investigated for specific examples of stocks, populations, or area-based changes. This work, however, would be more appropriate to specific assessors of these stocks, areas etc.

2.5.3 Task 2 – Document available cases in which biological or life-history parameter estimates were used as additional information to improve the understanding of the ICES/GFCM stock health

It is well-known that environmental conditions influence the development and health of individuals and populations, and this knowledge is considered from multiple perspectives and implemented within the advice process (from survey groups to benchmarks and stock assessments). A summary of case studies improving understanding of ICES/GFCM stock health is included in last year's report. As these case studies represent a reasonable summary of the state of knowledge necessary for inclusion in the broad remit of this Task, we consider last year's report sufficient for the time being. We therefore recommend that this Task should be closed.

2.5.4 Task 3 – Assess options to present biological parameters as supplementary diagnostics in addition to the standard graphs used in stock assessment and fisheries advice, within the scope of development towards ecosystem-based fisheries advice

The group agreed that currently body condition factor is certainly the most useful and urgent (additional) biological parameter to be considered when assessing health of fish stocks, in the Baltic and North Sea and the Mediterranean. From a stock assessment perspective, weight-at-age graphs should capture decreases in body condition. However, age comes with a bias and time-space coverage of weight data can be unbalanced which can mask the true changes in body factor and delay an early recognition of the increasing influence of environmental changes on stock health. The use of a body condition factor provides a more direct, unbiased view into stock health than weight-at-age data, although body condition cannot be directly used as an input parameter in stock assessment. One example of this is in the Eastern Baltic cod (Figure 2.5.4a), where interannual variation in body condition was monitored, with notable changes observed over time. This work highlighted the importance of routine compilation and inspection of additional biological parameters like condition factor, in which unexpected deviations should be detected early enough to take a closer look and understand the ecological mechanisms; so that ultimately appropriate management measures can be taken.

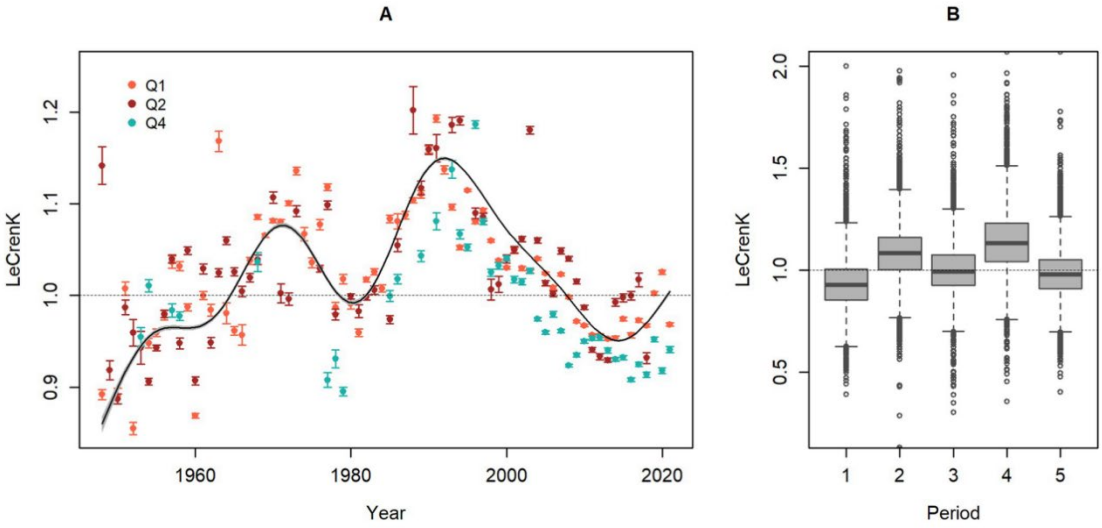

Fig 2. Time series of body condition of Eastern Baltic cod. (A) Developments in average LeCren's K condition index, by quarter (Q), but combined for all SDs and data sources (commercial and survey). The bars indicate standard error of the mean. The line shows smoothed year effects from GAM analyses including data source, Q and SD as categorical variables. (B) Variability in LeCren's K condition indices (Q1 and Q2 combined) in selected time periods representing peaks and troughs in the average condition: 1948–1955 (Period 1), 1970–1975 (Period2), 1978–1983 (Period 3); 1989–1993 (Period 4) and 2015–2021 (Period 5).

https://doi.org/10.1371/journal.pone.0286247.g002

Figure 2.5.4a. Example time-series of body condition in Eastern Baltic cod (from Eero *et al.*, 2023).

Similarly, spatial monitoring of condition factors can reveal key differences or changes in stock health, at scales relevant to benchmark and assessment work. An example of this for haddock in the North Sea is shown in Figure 2.5.2b.

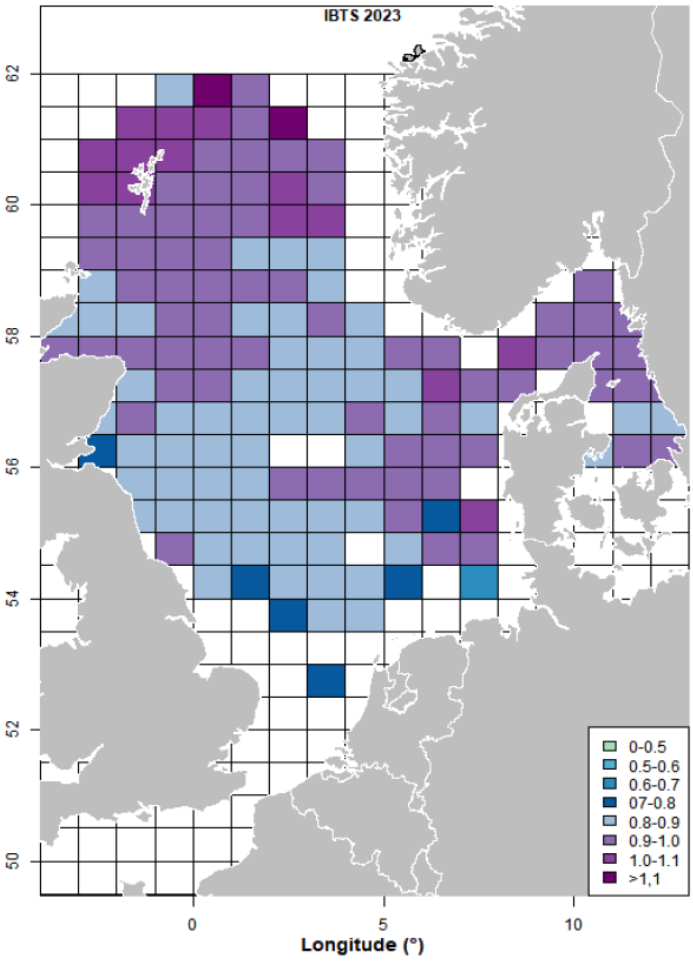

Figure 2.5.2b. Fulton's K condition of all the haddock weighed and measured individually during the NS-IBTS Q1 2023 (IBTSWG-report 2023).

It may therefore be more important to monitor and visualize the trends (temporal and spatial) of the additional biological and life-history parameters properly (e.g. on a yearly basis) rather than to necessarily focus on including these in the assessments.

The group considers the diagnostics and graphs given in last year's report as reasonable examples of possible options. Further examples would not change the message presented there, and the task is therefore considered final for the time being. We therefore suggest that this Task should be closed.

We do, however, recommend that **temporal trends in body condition factor be taken into consideration as routine diagnostics by the stock assessment teams both in the benchmarks and in the stock assessment reports**. Condition factor diagnostics graphs can be presented in different ways, for example the LeCrens condition index (e.g. Eero *et al.*, 2023) or length- or age-specific Fulton´s body condition factor, or separated by sex if relevant (e.g. in flatfishes). Mapping of condition factor distribution by ICES areas could also be considered to assess spatial patterns in the distribution of fish in lower or higher body condition. This would, however, require reasonable sampling coverage of individual length and weight measurements during the surveys (see Task 4 of quality assurance below).

Within the benchmarking and stock assessment processes, notable changes in body condition factor observed by the assessment group (relevant to each stock etc. being assessed) should trigger further investigation, e.g. on the weight-at-age data. These triggers and investigations should be case-specific (e.g. stock, population, area, etc.), but observation of temporal trends should be formalized within the process.

2.5.5 Task 4 – Assess accessibility of data and quality assurance of additional biological and life-history parameters

Accessibility of data: There are several sources to access data that can be used to derive additional biological parameters. These include:

- DATRAS and other international databases, e.g. to get length and weight data
- National data from pilot projects, BSc, MSc and PhD theses and other dedicated projects

Stock assessment teams could harvest these data sources and decide if additional data collection is necessary. If additional data are considered necessary, the assessment team should liaise with the respective survey working group.

Data quality assurance: The additional biological and life-history parameters we considered while working on this ToR E in recent years and in 2025 were mostly "simple" parameters based on length and weight measurements. We see the quality assurance of these parameters in the hands of the survey working groups, e.g. standardization of units (mm, cm, half cm etc.) for certain parameters, and use of calibrated balances, with appropriate precision both onboard the vessels and in the laboratories. However, special requests may come up and certainly be handled by WGBIOP but do not require a specific ToR or Task.

Summary of Tasks

1. We suggest that in benchmarks and assessment work, temporal (and spatial if relevant) trends in body condition factor (at scales relevant to the stock/area etc.) be included.
2. We suggest that further quality assurance be carried out by the survey groups to ensure maximum possible consistency between standard measured parameters such as fish length and fish weight.

3. We suggest that the conclusions of this ToR be circulated to each benchmark, stock assessment, and survey working group.
4. We suggest that Tasks 1-3 of ToR E be closed, and Task 4 be merged into ToR A in WGBIOP.

2.5.6 Workplan for 2026

Given the overly broad remit of the tasks, the progress achieved in the last years and the minimal need for updates to last year's report, except in cases where information on specific stocks or populations is requested, we suggest ToR E to be closed, while the quality assurance issue of additional biological parameters of Task 4 should be merged into ToR A. Thus, this ToR E could be closed, and the specific items integrated into another subgroup of WGBIOP.

2.6 ToR f. – Planning and creating reference collections and overseeing the maintenance of the reference collections

2.6.1 Progress during WGBIOP 2025

- Finalized the WGBIOP/WGSMART guide for TRAINING Reference Collections
- Screening for threshold values for AP, CV and APE, above which QC is considered acceptable
- Initiated discussions for defining the setup and application of QUALITY CONTROL Reference Collections

2.6.2 Task 1 – Liaise with WGSMART for development and feedback on the reference collections and training module

Liaison with WGSMART will be in the EU EMFAF project **Smart4SAM** (SmartDots for Monitoring, Accuracy and Reliable Training of essential biological data - Quality assurance for stock Assessment). Smart4SAM aims at developing, among other things, two new SmartDots modules for creating Reference Collections and for performing Quality Control. Both WGSMART and WGBIOP chairs are partners in the Smart4SAM project.

2.6.3 Task 2 – Provide guidelines for the setup of reference collections and their use for training, quality assurance and control

The following guidelines for the Introduction and for the TRAINING reference collection were discussed in detail and agreed upon. Some details will need to await specific outcomes of the Smart4SAM project. These are highlighted with comments and can be found on the WGBIOP SharePoint under
https://community.ices.dk/ExpertGroups/WGBIOP/SitePages/HomePage.aspx?RootFolder=%2FExpertGroups%2FWGBIOP%2F2025%20Meeting%20Documents%2F04%2E%20Working%20documents%2FToR%20F&FolderCTID=0x01200022642848DE2D35419F23C3B25CDCA7E2&View=%7BE2F6D274%2DDF6B%2D4CFC%2DBA15%2D3D78059782BA%7D

2.6.4 Task 3 – Compile a list of existing reference collections, where to find them and who to contact

This task was completed during WGBIOP 2024. A list of existing TRAINING reference collections including their status was compiled, including an overview of potential SmartDots samples from earlier exchanges that could be used. This overview may be found in the 2024 report.

2.6.5 Task 4 – Integrate existing reference collections in SmartDots once the new module is available

A decision has been made to use events in SmartDots as the primary reference collections. If additional reference data are needed for specific species or stocks, users are encouraged to add such collections to SmartDots in accordance with the established guidelines for reference collections. Examples of potential additions include known-age eel data from Sweden and Baltic herring data from Finland and Sweden. Questionnaire Annex 4 from WGBIOP report from 2022 was revisited for any precise figures used in quality assurance or quality checks, but as the questionnaire was not aimed to collect answers to any specific QA/QC thresholds none was found. If such data are needed a new query for it should be sent to age coordinators and to be analysed in WGBIOP 2026.

2.6.6 Task 5 – Supervise the integration of new samples in the reference collections by the event coordinators

TRAINING RC: We have identified criteria for which samples of existing SmartDots events are suitable for being included in RCs (see guidelines). Those will be identified with an extra column in the metadata, where event coordinators indicate for each sample, whether it is suitable or not.

2.6.7 ToR f workplan for 2026

Task 1 Liaise with WGSMART for development and feedback on the reference collections and training module

A summary of the Smart4SAM progress will be provided in 2026.

Task 2 – Provide guidelines for the setup of reference collections and their use for training, quality assurance and control

The following tasks remain to be addressed:

TRAINING RC:

- Follow up on precision thresholds for calibration events (PGCCDBS Annex 4 + SLU protocol) and summarize these into meaningful guidelines.
- Identify the reason why the CVs in all SmartDots events are so high, even when PA is very high. Virtually no event has a CV < 5% as recommended by EFAN.

QUALITY CONTROL RC:

- Intersessional work, mainly within Smart4SAM to design sample selection and analytical approaches for QC protocols in collaboration with WGSMART
- Contact Silvia O'Sullivan, Quality Control responsible at NOAA, for protocols and possibly presentation at WGBIOP 2026.

Task 3 – Compile a list of existing reference collections, where to find them and who to contact

No further work planned.

Task 4 – Integrate existing reference collections in SmartDots once the new module is available

No further work planned.

Task 5 – Supervise the integration of new samples in the reference collections by the event coordinators

Will be dealt with in Smart4SAM.

3 Other achievements

The 2025 WGBIOP meeting achievements included the WKETAC workshop, the Smart4SAM project and scientific presentations on *Solea solea* that combined multidisciplinary analyses of otolith and ovary data. Efforts were made to liaise with other Working Groups to strengthen collaboration and to create a GitHub page containing key documents that may also be valuable to other WGs. Together, these achievements highlight WGBIOP's input in methodological innovation and practical advancements in biological data for sustainable fisheries advice.

3.1 WKETAC

The 2025 WKETAC workshop marked a key scientific achievement for WGBIOP, bringing together fisheries biologists and computational experts to evaluate the latest automated technologies—such as deep learning, computer vision, and two-dimensional/three-dimensional imaging—for extracting age, life history, and stock identification data from calcified structures like otoliths and scales. The workshop's outcomes directly advance WGBIOP's mission of improving the accuracy, efficiency, and reproducibility of biological parameters by rigorously assessing the cost–benefit, robustness, and suitability of emerging automated methods before they are adopted in stock assessment and data collection frameworks central to WGBIOP's remit. By identifying priority species for pilot applications and developing best-practice guidelines for transparent and reproducible implementation, WKETAC strengthens quality assurance and fosters scientific innovation across ICES and partner institutions, positioning WGBIOP at the forefront of integrating next-generation approaches into fisheries science and advice.

3.2 Smart4SAM

The Smart4SAM presentation showcased a major scientific advance within WGBIOP's remit—delivering an integrated management framework for the quality assurance and operational use of age and maturity data in fish stock assessment. Smart4SAM expands the ICES SmartDots platform by allowing creation of public reference collections, self-training modules, robust quality control procedures, and the seamless quantification and integration of ageing uncertainty in assessment models, addressing long-standing priorities identified by WGBIOP. The project brings together partners across Europe to standardize and validate new AI-assisted approaches, deliver user-friendly tools for laboratories, improve training and technician performance, and ensure transparent communication across the research and assessment communities, thus positioning WGBIOP at the forefront of modernising biological parameter workflows and supporting the reliability of scientific advice.

3.3 Scientific highlight: advancing life-history biological parameters for stock assessment and ecosystem-based fisheries management through otolith and ovary analyses: the sole case

The scientific session on sole brought together complementary perspectives on the value of calcified structures and gonad analyses for understanding this key demersal species.

3.3.1 "What do we expect to learn from otolith analyses? A Sole example" (Audrey Geffen)

Audrey Geffen's presentation, "What do we expect to learn from otolith analyses? A Sole example," highlighted the value and realistic expectations of using diverse biomarkers from otoliths and other tissues for assessing fish age, growth, condition, and population structure. By synthesizing multidisciplinary markers—genetic, environmental, and phenotypic—her work demonstrates a powerful integrative approach to understanding stock identity, migration, and life-history traits in *Solea solea*. The presentation emphasized the need for coordinated sampling design and multi-marker analysis to effectively support fisheries management and traceability, aligning closely with WGBIOP's goals of improving biological parameter data quality and ecological interpretation.

3.3.2 "Insights into the life history of sole (*Solea solea*) from ovary and otolith data" (Tuan-Anh Bui)

This was followed by Tuan Anh Bui's presentation, "Insights into the life history of sole (*Solea solea*) from ovary and otolith data," which provided important findings on reproductive investment and growth dynamics in sole populations across multiple North Sea ecosystems. Using extensive ovary weight and maturity stage data alongside otolith growth increments, the study revealed how reproductive effort scales with environmental temperature and body size, showing hyper-allometric reproductive investment and differential growth responses to warming. These insights contribute vital biological parameters that underpin stock assessments and fisheries advice, directly supporting WGBIOP's mandate to refine and validate key life-history data to improve fisheries management outcomes.

3.4 Liaise with other working groups

Numerous Working Groups were invited to present the work they have carried out in recent years and to discuss any issues relevant to WGBIOP. The aim of these exchanges was to strengthen collaboration and streamline activities. Presentations were given by WGRDBESGOV, WGBIOPTIM, WGSMART, RCG North Sea and Baltic, and DATRAS. All presentations from these groups are available on the WGBIOP SharePoint.

GitHub has become an essential platform within WGBIOP, serving as a central repository for shared code, scripts, and documentation related to biological parameter estimation. By maintaining open repositories, WGBIOP promotes transparency and reproducibility, allowing clear tracking of changes and fostering confidence in analytical outputs. This platform is also useful for collaborating with Working Groups—such as survey and assessment groups.
The following documents are currently available:

- Guidelines for age reading exchanges and workshops
- Guidelines for maturity exchanges and workshops
- Maturity leaflet
- Maturity conversion tables between maturity scales
- Roles of age readers and maturity coordinators
- Stock-specific spawning periods

4 Next meeting

It was agreed that the next meeting would be hybrid. The meeting is scheduled from 5 to 9 October 2026, in Nea Peramos, Kavala, Greece, hosted by the Greek Fisheries Research Institute (INALE – ELGO Dimitra). The hybrid approach will allow for greater flexibility and more active participation, ensuring that all members can contribute effectively to WGBIOP's ongoing work.

5 Reference list

Andrialovanirina N., Hache À., Mahe K., Couette S., Poisson Caillault E. (2022). Data applied to automatic method to transform routine otolith images for a standardized otolith database using R. SEANOE, https://doi.org/10.17882/91023

Andrialovanirina N., Poloni L., Laffont R., Poisson Caillault E., Couette S., Mahé K. (2024). 3D otolith mesh dataset of red mullet (Mullus barbatus) in the Mediterranean Sea. SEANOE, https://doi.org/10.17882/99980

Andrialovanirina N., Poisson Caillault E., Mateos A., Laffont R., Mackenzie K., Couette S., Mahe K. (2025). 3D Meshes Dataset of Sagittal Otoliths from 22 Fish Species in the English Channel and North Sea. SEANOE, https://doi.org/10.17882/106566

Campana, S. E. 2001. Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. Journal of Fish Biology, 59 (2), 197–242 | https://doi.org/10.1111/j.1095-8649.2001.tb00127.x

Cayetano A., Stransky C., Birk A., Brey T., (2024). Fish age reading using deep learning methods for object-detection and segmentation. ICES Journal of Marine Science, 81(4), 687–700, https://doi.org/10.1093/icesjms/fsae020

Eero M, Brander K, Baranova T, Krumme U, Radtke K, Behrens JW (2023) New insights into the recent collapse of Eastern Baltic cod from historical data on stock health. PLoS One 18(5):e0286247, https://doi.org/10.1371/journal.pone.0286247

ICES 2011. Report of the Workshop of National Age Readings Coordinators (WKNARC), 5-9 September 2011, Boulogne-sur-Mer, France. ICES CM 2011/ACOM:45. 175 pp. https://doi.org/10.17895/ices.pub.19280897

ICES. 2014. Report of the Workshop for maturity staging chairs (WKMATCH), 11–15 June 2012, Split, Croatia. ICES CM 2012/ACOM:58. 57 pp.

ICES, 2023. International Bottom Trawl Survey Working Group (IBTSWG). ICES Scientific Reports. 5:80. 204 pp. https://doi.org/10.17895/ices.pub.23743989

Katija, K., Orenstein, E., Schlining, B., Lundsten, L., Barnard, K., Sainz, G., ... and Bell, K. L. (2022). FathomNet: A global image database for enabling artificial intelligence in the ocean. Scientific reports, 12(1), 15914

Jan, M., Spangaro, A., Lenartowicz, M., and Mattiazzi Usaj, M. (2024). From pixels to insights: Machine learning and deep learning for bioimage analysis. BioEssays, 46(2), 2300114

Lombarte, A., . Chic, V. Parisi-Baradad, R. Olivella, J. Piera and E. Garca-Ladona. (2006). A web-based environment from shape analysis of fish otoliths. The AFORO database. Scientia Marina, 70, 147-152, http://aforo.cmima.csic.es/

Smoliński, S., Schade, F. M., and Berg, F. (2020). Assessing the performance of statistical classifiers to discriminate fish stocks using Fourier analysis of otolith shape. Canadian Journal of Fisheries and Aquatic Sciences, 77(4), 674-683

Miyajima-Taga, Y., Suzuki, K., Takahashi, M., Hirabayashi, Y., Oi, K., and Inoue, N. (2025). Importance of input image size on the performance of automatic age determination of chum salmon Oncorhynchus keta using deep learning. Fisheries Science, 1-16

Annex 1: List of participants

Name	Institute	E-mail	Country (of institute)
Andrés Villaverde	IEO-CSIC	andres.villaverde@ieo.csic.es	Spain
Ângela ML. Canha	Regional Directorate for Fisheries	angela.ml.canha@azores.gov.pt	Portugal
Annelie Hilvarsson	SLU Department of Aquatic Resources	annelie.hilvarsson@slu.se	Sweden
Cindy van Damme	Wageningen University and Research	cindy.vandamme@wur.nl	Netherlands
Côme Denechaud	Institute of Marine Research	come.denechaud@hi.no	Norway
Cristina Bultó	IEO-CSIC	cristina.bulto@ieo.csic.es	Spain
Cristina Nunes	Portuguese Institute for the Sea and the Atmosphere	cnunes@ipma.pt	Portugal
David Maxwell	Cefas Lowestoft Laboratory	david.maxwell@cefas.gov.uk	UK
Dolores Garabana	IEO-CSIC	dolores.garabana@ieo.csic.es	Spain
Elor Sepp	Estonian Marine Institute	elor.sepp@ut.ee	Estonia
Francesca Vitale	SLU Department of Aquatic Resources Institute of Marine Research	francesca.vitale@slu.se	Sweden
Gráinne Ní Chonchúir	Marine Institute	grainne.nichonchuir@marine.ie	Ireland
Gualtiero Basilone	National Research Council	gualtiero.basilone@cnr.it	Italy
Guðrún Finnbogadóttir	Marine and Freshwater Research Institute	gudrun.finnbogadottir@hafogvatn.is	Iceland
Inês Farias	Portuguese Institute for the Sea and the Atmosphere	ifarias@ipma.pt	Portugal
Jari Raitaniemi	Natural Resources Institute Finland - Turku	jari.raitaniemi@luke.fi	Finland
Joana Ribeiro	International Council for the Exploration of the Sea	joana.ribeiro@ices.dk	International
Joanne Smith	Centre for Environment, Fisheries and Aquaculture Science	joanne.smith@cefas.gov.uk	UK
Julie Coad Davies	DTU Aqua, National Institute of Aquatic Resources	joco@aqua.dtu.dk	Denmark

Name	Institute	E-mail	Country (of institute)
Karen Bekaert	Flanders Research Institute for Agriculture, Fisheries and Food	karen.bekaert@ilvo.vlaanderen.be	Belgium
Karin Hüssy	DTU Aqua, National Institute of Aquatic Resources	kh@aqua.dtu.dk	Denmark
Kélig Mahé	Ifremer, Centre de Manche Mer du Nord	kelig.mahe@ifremer.fr	France
Kirsteen MacKenzie	Ifremer, Centre de Manche Mer du Nord	kirsteen.mackenzie@ifremer.fr	France
Konstantina Ofridopoulou	Aristotle University of Thessaloniki and Fisheries Research Institute – ELGO Dimitra	kofry@bio.auth.gr, ofridopoulouk@inale.gr	Greece
Larissa Modica	IEO-CSIC Centro Oceanográfico de Santander	larissa.modica@ieo.csic.es	Spain
Laurent Dubroca	Ifremer, Centre de Manche Mer du Nord	laurent.dubroca@ifremer.fr	France
Léa Joly	Centre de Manche Mer du Nord	Lea.Joly@ifremer.fr	France
Lukas Snoeck	Flanders Research Institute for Agriculture, Fisheries and Food	Lukas.Snoeck@ilvo.vlaanderen.be	Belgium
Mandy Gault	Marine Laboratory	mandy.gault@gov.scot	UK
Maria Krüger-Johnsen	DTU Aqua, National Institute of Aquatic Resources	mkru@aqua.dtu.dk	Denmark
Michelle Douglas	Scottish Government, Marine Laboratory	michelle.douglas@gov.scot	UK
Nicolas Andrialovanirina	Ifremer, Centre de Manche Mer du Nord	nicolasandrialova@gmail.com	France
Patricia Cortegoso	IEO-CSIC	patricia.cortegoso@ieo.csic.es	Spain
Patrícia Gonçalves	Portuguese Institute for the Sea and the Atmosphere	patricia@ipma.pt	Portugal
Pedro Torres	IEO-CSIC Centro Oceanografico de Málaga	pedro.torres@ieo.csic.es	Spain
Pierluigi Carbonara	COISPA Tecnologia and Ricerca	carbonara@coispa.it	Italy
Ralf van Hal	Wageningen Marine Research	ralf.vanhal@wur.nl	Netherlands
Remigiusz Szymański	National Marine Fisheries Research Institute	rszymanski@mir.gdynia.pl	Poland
Renato Cruz	Regional Directorate for Fisheries	renato.es.cruz@azores.gov.pt	Portugal

Name	Institute	E-mail	Country (of institute)
Romain Elleboode	Ifremer, Centre de Manche Mer du Nord	romain.elleboode@ifremer.fr	France
Rosalia Ferrerri	National Research Council	rosalia.ferreri@cnr.it	Italy
Vesala Sami	Natural Resources Institute Finland (Luke)	sami.vesala@luke.fi	Finland
Sarah Johnstone	Marine Lab, Aberdeen	sarah.johnstone@gov.scot	UK
Szymon Smoliński	National Marine Fisheries Research Institute	ssmolinski@mir.gdynia.pl	Poland
Uwe Krumme	Thünen-Institute of Baltic Sea Fisheries	uwe.krumme@thuenen.de	Germany
Valerio Visconti	Cefas Lowestoft Laboratory	valerio.visconti@cefas.gov.uk	UK
Vasiliki Kousteni	Fisheries Research Institute – ELGO Dimitra	kousteni@inale.gr	Greece
Yvette Heimbrand	SLU Department of Aquatic Resources	yvette.heimbrand@slu.se	Sweden

Annex 2: Resolutions

Term 2/3

2023/MT/DSTSG05

The **Working Group on Biological Parameters** (WGBIOP), chaired by Karen Bekaert, Belgium, Konstantina Ofridopoulou, Greece, and Valerio Visconti, UK, will work on ToRs and generate deliverables as listed in the Table below.

	Meeting dates	Venue	Reporting details
Year 2024	7–10 October 2024	Palermo, Italy	E-evaluation and interim report by 20 December 2024 to DSTSG
Year 2025	6–10 October 2025	Boulogne-sur-Mer, France	E-evaluation and interim report by 19 December 2025 to DSTSG
Year 2026	TBD 2026	TBD Kavala, Greece or Iceland	E-evaluation and final report by 18 December 2026 to DSTSG

ToR descriptors

ToR	Description	Background	Science plan codes	Duration	Expected deliverables
a	Quality assurance.	Organising, reviewing and prioritizing exchanges and workshops requested from EGs, WKs and other ICES-related groups represented the base for the quality assurance routine of this ToR. It is also vital to plan these in line with the ICES benchmark schedule and boost the communication loop with assessment groups. Also, updating and maintaining guidelines is a key point for quality assurance.	3.1, 3.2	Year 1, 2, 3	1 Coordinate communication between WGBIOP members and the corresponding stock assessment coordinator/group of interest based on the current rolling issues list and/or other emerging issues and recommendations. 2 Complete annual overview of planned studies, ex-changes and workshops (i.e. update of master tables) and planning of new calibration events. 3 Review and update the guidelines for calibration events and reports to be published in the ICES Library including fixing DOI where applicable (e.g. SmartDots publications). 4 Overview and recommendation of validation studies (age, maturity, etc.).

b	Promoting the unequivocal understanding and adoption of the SMSF scale as well as advising on the histological approach as a validation method of maturity staging.	Since 2020, only a few countries have reported maturity using the SMSF scale, and several institutes and WGs pointed out confusion when it comes to the interpretation of the (sub)stages. Given that histology is the key to accurate maturity staging, planning its routine use is the goal. Also, the participation of maturity experts within WGBIOP is currently limited and expanding this expertise basis is desirable.	3.1, 4.1	Year 1, 2, 3	1 Liaise with WGs to clarify the (sub)stages, ensuring the proper adoption of SMSF scale and receiving feedback on potential issues related to maturity. 2 Produce stock-specific tables with the identification of the main timing of gonadal development and spawning period supporting the proper use of the SMSF scale. 3 Correct the previously published conversion tables to the SMSF scale. 4 Draft a working plan for the adoption of the histological approach as a validation or estimation method, and liaise with RCGs to establish the working plan at the European scale according to the evaluation needs. 5 Encourage evenly distributed WGBIOP participation of experts on maturity and other biological parameters.
c	Follow-up on emerging tools and methods for the determination of biological parameters.	Many new methods for the determination of biological parameters, such as Artificial Intelligence (AI), genetics, shape analysis, otolith microchemistry, etc., have been developed over the last years. This is of interest to WGBIOP who aims to be aware of the latest developments. The goal of this ToR is to monitor, coordinate and facilitate these developments where possible.	3.1, 4.1, 4.3	Year 1, 2, 3	1 Liaise with relevant ICES Working Groups (e.g. WGMLEARN, WGSMART, SIMWG). 2 Facilitate the development of emerging tools related to biological parameters (e.g. set up a central image repository for AI). 3 Provide guidelines for standardization of methods and protocols for emerging tools.

d	Review the procedures used in calibration events and facilitate the transfer of error data into stock assessments.	The statistical methods applied to analyse the results from SmartDots calibration events haven't been revised since 2000. In recent years, progress has been made by WGBIOP, in cooperation with WGSMART, to identify errors in age and maturity estimations via calibration events. Steps towards the incorporation of age and maturity estimation errors into the stock assessment process have been made (WKMACQI, 2018; WKAMEMSA, 2021), and some stock assessment and benchmark groups have incorporated this error data into the stock assessment model runs (WKBALTPEL and WKSSNSK). Further work is required to correctly identify and account for the error sources included in these events. The aim of this ToR is to revise the statistical methods used to identify the errors in calibration events, as well as improve the sampling design for setting up calibration events and facilitating the integration of error data in stock assessments.	3.1	Year 1, 2, 3	1 Liaise with relevant assessment working groups to facilitate the use of age error data in stock assessment. 2 Revise statistical procedures to improve the reporting of calibration events, e.g. by taking into account the age plus groups, and quality scores (AQ or QS) and investigate temporal and spatial stratification of samples at national and/or regional level. 3 Identify potential sources of errors in the procedures involved in generating Error Matrices (EM) or the raw data supplied to the stock assessors. 4 Develop an improved sampling design for SmartDots calibration events to account for these sources of error. 5 Organize a second workshop on the use of Ageing and Maturity Staging Error Matrices in Stock Assessment (WKAMEMSA).
e	Potential role of additional biological and life-history parameters in stock assessments and fisheries advice and evaluation of significant changes in time.	The scope of this ToR is to assess the possibility of improving the quantity and quality of the data used in assessment and advice, considering data availability.	3.2, 4.1, 5.1	Year 1, 2, 3	1 Review the links between biological and life-history parameters and (climate-induced) changes in environmental conditions. 2 Document available cases in which biological or life-history parameter estimates were used as additional information to improve the understanding of the ICES/GFCM stock health. 3 Assess options to present biological

					parameters as supplementary diagnostics in addition to the standard graphs used in stock assessment and fisheries advice, within the scope of development towards ecosystem-based fisheries advice. 4 Assess the accessibility of data and the quality assurance of additional biological and life-history parameters.
f	Planning and creating reference collections and overseeing the maintenance of the reference collections.	The scope of this ToR is to develop an approach for the creation of reference collections. A reference collection is a collection of images of a validated biological parameter (e.g. maturity) or consensus-derived results (e.g. age). The purpose of a reference collection is to have a set of reference materials for calibration and training of new and established readers. It is important, therefore, that reference sets are suitable for calibration and training purposes. Consistency in approach, together with predefined requirements, is key.	3.1, 4.1	Year 1, 2, 3	1 Liaise with WGSMART for development and feedback on the reference collections and training module. 2 Provide guidelines for the setup of reference collections and their use for training, quality assurance and control 3 Compile a list of existing reference collections, where to find them and whom to contact 4 Integrate existing reference collections in SmartDots once the new module is available. 5 Supervise the integration of new samples in the reference collections by the event coordinators

Summary of the work plan

Year 1	Evaluate the quality of biological parameters used in assessments and coordinate communication with the corresponding assessment coordinators. Promote the correct use and adoption of the new scale of maturity stages. Monitor, coordinate and facilitate the development of emerging tools concerning the biological parameters. Review the procedures used in calibration events and their error mitigation. Investigate the possibility of improving the quantity and quality of the data used in assessment and advice according to data availability. Scheduling of exchanges, workshops and validation studies aligned with the benchmark cycle, and creation and maintenance of reference collections and guidelines.
Year 2	Evaluate the quality of biological parameters used in assessments and coordinate communication with the corresponding assessment coordinators. Promote the correct use and adoption of the new scale of maturity stages. Monitor, coordinate and facilitate the development of emerging tools concerning the biological parameters. Review the procedures used in calibration events and their error mitigation. Investigate the possibility of improving the quantity and quality of the data used in assessment and advice according to data availability. Scheduling of exchanges, workshops and validation studies aligned with the benchmark cycle, and creation and maintenance of reference collections and guidelines.
Year 3	Reviewing the status of issues, achievements and developments concerning biological parameters and quality assurance of life-history parameters provided for assessment and management processes. Reviewing the emerging tools and database developments for providing and accessing biological parameters information. Identify future needs in line with the ICES objectives and Science Plan and the wider marine environmental monitoring and management within Europe and propose a future work plan improving quality assurance of biological parameters.

Supporting information

Priority	The main objective of WGBIOP will be to support the development and quality assurance of regional and national provision of biological parameters as reliable input data to integrated ecosystem stock assessment and advice while making the most efficient use of expert resources. As biological parameters are among the main input data for most stock assessments and mixed fishery modelling, these activities are considered to have a very high priority. All National Age Reader/Maturity Stager Coordinators (ICES and GFCM) will be invited. Experts relevant to the current benchmarks of the year of WGBIOP will be invited as well as relevant external experts such as statisticians or specific EG members.
Resource requirements	None.
Linkages to ICES committees or groups	There is a very close working relationship with all the groups in DSTSG, as well as BOG and ACOM. It is also very relevant to the WGMLEARN, WGSMART, SIMWG.
Linkages to other organizations	GFCM.

Annex 3: Exchanges and workshops (ToR a)

Exchanges and workshops completed in 2024 and 2025

North East Atlantic Mackerel (*Scomber scombrus*] Age Reading Exchange (SmartDots event ID 1888), coordinated by Gráinne Ní Chonchúir (Marine Institute)

1. mac.27.nea - Mackerel (*Scomber scombrus*) in subareas 1-8 and 14 and in divisions 9.a, 12.a and 12.b North East Atlantic and adjacent waters. This stock is a category 1, age-based assessment and comes under the umbrella of the Working Group on widely distributed stocks (WGWIDE). Northeast Atlantic Mackerel was the subject of a benchmark review in March 2025, and the outputs of this exchange were sent to the relevant stock coordinator in time for consideration at the benchmark meeting. Previous exchanges and workshops resulted in relatively low overall agreement, and this combined with the upcoming benchmark was the motivation to conduct another calibration exercise.
2. A total of 268 sectioned otolith images were used for this exchange, which span the full Spatial/temporal and length/age ranges of the stock.

   15 Readers from 9 Member States participated in this exchange, consisting of 9 Advanced readers – whose ages are used in the assessments and 6 Basic readers. The list of participants can be found below.

Participants List and Expertise

Reader code	Laboratory	Country	Expertise	Readings In Assessment
R02 IS	MFRI	Iceland	Advanced	Yes
R04 IE	MI	Ireland	Advanced	Yes
R06 NL	WMR	Netherlands	Advanced	Yes
R08 DK	DTU-Aqua	Denmark	Advanced	Yes
R10 IE	MI	Ireland	Advanced	Yes
R12 ES	IEO	Spain	Advanced	Yes
R14 FO	FMRI	Faroe Islands	Advanced	Yes
R16 FO	FMRI	Faroe Islands	Advanced	Yes
R18 NL	WMR	Netherlands	Advanced	Yes
R24 FR	Ifremer	France	Basic	No
R30 PT	IPMA	Portugal	Basic	No
R32 DK	DTU-Aqua	Denmark	Basic	No

R34 DE	Thünen-Institut	Germany	Basic	No
R40 ES	AZTI	Spain	Basic	No
R42 ES	AZTI	Spain	Basic	No

Table 4.1 Spatial and temporal distribution of samples used in this exchange

Component	ICES Area	Subarea	N Images			Length Range (cm)	Image Origin
			Sem 1	Sem 2	Total		
South	8c	8cE	5	5	10	29 - 43	IEO Spain
		8cW	5	5	10		
	9a	9aN	5	5	10	20 - 42	
		9aCN	14	13	27		IPMA (Portugal)
		9aC					
		9aCS					
		9aS					
West	6	6a	3	7	10	27 - 37	WMR (The Netherlands)
			18	4	22	35 - 40	Marine Institute (Ireland)
	7	7b	5		5	31 - 39	IMR (Norway)
		7c	5		5	34 - 38	TI - SF (Germany)
		7g	5		5	19 - 33	
		7d		5	5	26 - 35	
		7f		5	5	21 - 29	IFREMER (France)
		7h		5	5	16 - 30	
		7j		5	5	17 - 33	
	8abde	8a		10	10	19 - 30	
		8b	10		10	32 - 39	AZTI (Spain)
North Sea	4	4a	5	15	20	19 - 43	TI - SF (Germany)
		4b	6	14	20	19 - 39	WMR (The
				2	2	19	DTU - Aqua
		4c	5		5	19 - 24	WMR (The Netherlands)
				7	7	15 - 20	DTU - Aqua
North Distribution	2	2a	20		20	16 - 39	DTU - Aqua
		2b		20	20	33 - 42	IMR (Norway)
	5	5b	10		30	26 - 40	
			7	13			FMRI (Faroe Islands)
	Total		128	140	268		

1. Results:
    a. PA, CV, APE for advanced and all readers separately (AQ/QS 1 and 2 only).

Summary of statistics; Total number of samples (NSample), coefficient of variance (CV), percentage of agreement (PA) and average percentage error (APE) for **all ages and readers**

NSample	CV	PA	APE
268	26%	69%	14%

Summary of statistics; Total number of samples (NSample), coefficient of variance (CV), percentage of agreement (PA) and average percentage error (APE) for all ages for **Advanced Readers** only.

NSample	CV	PA	APE
268	21%	73%	11%

Results from advanced readers show an overall agreement of 73%, an improvement from 69% when less experienced readers were included and up from 67.8% in the previous exchange in 2021. Agreement is highest at age 0 (96%) and remains strong for ages 1-6 (71-82%), while it is lowest for ages above 10 (37-49%). The coefficient of variation (CV) is highest for ages 1-4 (up to 38%) and stabilizes for ages 6-12 (6-17%), with some increased variability at older ages. The percentage agreement varies by reader, however overall, there is no strong systemic bias, but inconsistencies are notable at specific ages and among certain readers.

For basic readers, the overall agreement is lower compared to advanced readers, as expected, indicating greater variability of age readings. Basic readers tend to have higher disagreement, particularly at older ages where precision is more challenging. The coefficient of variation (CV) is also generally higher, reflecting inconsistencies in age estimation. Some readers exhibit noticeable biases, either underestimating or overestimating ages, which impacts overall reliability.

It's important to note that Basic readers tend to be new to age reading or to reading this particular species and are still training, which are factors that may contribute to the observed discrepancies.

While a marked improvement is observed, when compared with the previous exchange results in 2021, there is still room for improvement across all readers. Specific issues for individual readers to work on, are highlighted in the report.

The exchange findings highlight priorities for future work, i.e. the creation of a Training Reference Collection, active engagement with technological developments around machine learning, through WGBIOP workshops and WGSMART. Additional training and calibration to improve consistency and accuracy in age estimations is also recommended.

2. Conclusions, further actions needed:

Overall, the data highlights improved agreement, reduced variability of key age groups, and minimal systemic bias, across both Advanced and Basic readers which are all positive outcomes for the accuracy and reliability of mackerel age determination.

Age determination accuracy improved, particularly in younger age groups, but the results also underscore the challenges associated with older age classes. Future efforts should focus on refining methodologies for age determination in older fish, addressing inconsistencies among readers.

It is recommended to create a Training Reference Collection (TRC) following the guidelines proposed by WGBIOP in its 2024 Working Group report which will be available: https://www.ices.dk/community/groups/pages/wgbiop.aspx.

This is a very worthwhile exercise and will be a useful training tool for age readers. It may be possible to use a reference collection already compiled by Cefas as the basis for this.

It is recommended to actively engage and collaborate with WGBIOP and other Expert Groups e.g. The Working Group on Machine Learning in Marine Science (WGMLEARN) which reviews and identifies machine learning (ML) applications in marine science as well as emerging developments in the field of AI and ML. This group aims to identify key challenges and provide guidance and resources for data sharing (methods, training sets, protocols). Discussions around the fields of computer vision, Near-infrared spectroscopy, multidimensional shape analysis and otolith microchemistry and attempts at developing automated methods to estimate individual age and maturity will be of particular interest to age readers. https://www.ices.dk/community/groups/Pages/WGMLEARN.aspx

WGBIOP has also developed a draft proposal for a Workshop on Emerging Technologies for the Automated analysis of Calcified structures (WKETAC), possibly for 2025 which will also be of interest and would benefit from the active engagement of age readers.

Another exchange should be completed, after progress has been made on the actions recommended above, in order to ascertain if they have positively impacted age estimation for mackerel.

Northern and central Bay of Biscay common sole 27.8ab (*Solea solea*) Otolith Exchange **[(SmartDots event 1903)]** coordinated by Romain Elleboode (Ifremer)

1. This otolith exchange was performed on common sole (*Solea solea*, Aphia code 127160), in Subdivisions 8ab (Northern and central Bay of Biscay) for stock sol.27.8ab, and was coordinated by Romain Elleboode and Kirsteen MacKenzie. This stock is in category 1, with an age-based assessment managed by WGBIE assessment working group. The background and aim of the exchange were to intercalibrate the age readings and be able to identify possible issues as part of a European exchange.

2. A total of 310 otoliths were used for the exchange. The preparation methods used were both sectioned (SE), and sectioned and stained (SS) otoliths, and all photos were taken with reflected light. Eight age readers from five countries participated, with half being advanced readers involved in stock assessments. All samples were collected in 2023 and 2024, with the majority of samples originating in ICES areas 27.8b (only 47 samples from ICES area 27.8.a were used). Samples represented all four quarters (except for age 0, which were only present in Q4). The list of participants can be found below.

Participants List and expertise

Country	Expertise
Belgium	Advanced (SS)
Greece	Basic
Italy	Basic
France	Advanced (S)

France	Advanced (S)
Denmark	Basic
Belgium	Advanced (SS)
Italy	Basic

3. Results:

Ages ranged from 0 to 21 years. The results differed between reader categories, with an overall PA of 73% (80% for only advanced readers), a CV of 20% (16% for only advanced readers), and an APE of 12% for all readers (8% for only advanced readers). Results also differed between preparation methods, with a higher precision in age readings for images of sectioned otoliths than for images of sectioned and stained otoliths, despite a lower age range in SS otoliths. When considering only expert readers for each method, there was a lower coefficient of variation (CV= 13% SE, 11% SS) and high agreement among readers (PA= 87% SE, 82% SS). A decline in reading accuracy was noted from age 8 onwards.

4. Conclusions, recommendations, further actions needed:

Readers working regularly on stained samples read unstained samples with relative ease, but readers working on unstained samples found stained samples more difficult to interpret, leading to lower overall PA for SS samples. The main age reading challenges identified were the presence of false rings, ambiguities in edge interpretation, and clipped edges in older otoliths. Preparation methods should be considered carefully for future age reading on common sole.

North Sea, eastern English Channel whiting 27.47d (*Merlangius merlangus*) Otolith exchange **(SmartDots event 2941**), coordinated by Romain Elleboode (Ifremer) and Karen Bekaert (ILVO)

1. This otolith exchange was performed on whiting (*Merlangius merlangus*, Aphia code 400587), in Subdivisions 7d, 4c, 4b (North Sea, eastern English Channel) for stock whg.27.47d, and was coordinated by Karen Bekaert and Romain Elleboode. This stock is in category 1, with an age-based assessment managed by WGNSSK assessment working group. The background and aim of the exchange were to intercalibrate the age readings and be able to identify possible issues as part of a European exchange. A total of 424 images were obtained, including 97 sectioned (SE) otoliths (transmitted (ST) and reflected light (SR)), 50 French sectioned (SE) otoliths (transmitted (ST) light only, with MERLMNG in the name), and 90 broken otoliths (BR; transmitted (TLX) and reflected light (RLX)).

2. Sixteen age readers from eight countries participated, with half being advanced readers involved in stock assessments. All samples were collected between 2020 to 2025 A balanced distribution of samples was achieved for the sectioned otoliths across zones 27.7d, 27.4cb. The broken otoliths were mainly distributed in the North Sea, particularly in areas 27.4c-a. Samples represented all four quarters. *(except for age 0, which were only present in Q4).* The list of participants can be found below.

List of participants and expertise

Country	Expertise
Belgium	Basic
UK	Basic
Netherlands	Basic
France	Basic
France	Advanced
Germany	Advanced
Norway	Basic
Denmark	Advanced
Belgium	Advanced
UK	Basic
Denmark	Advanced
Netherlands	Basic
UK	Basic
Norway	Advanced

Germany	Advanced
UK	Advanced

3. Results:

Ages in the exchange ranged from 1 to 7 years. When considering **all** readers and all otolith samples, the weighted average percentage agreement (PA) based on modal ages was 63%, with a weighted average CV of 30% and an APE of 21%. Restricting the analysis to **sectioned** otoliths and the readers who assessed them produced identical summary statistics (PA = 63%, CV = 30%, APE = 21%). Similarly, the results for **broken** otoliths and their corresponding readers were nearly the same (PA = 63%, CV = 30%, APE = 20%).

Across **advanced** readers, the overall agreement metrics were a PA of 65%, a CV of 30%, and an APE of 21%. The minimal difference between advanced readers and all readers is noteworthy and reflects that some advanced readers performed less consistently than expected, diminishing the contrast between reader categories. **Broken** otoliths performed better than **sectioned** otoliths, with higher percentage agreement (70% vs. 63%) and lower CV (25% vs. 31%) and APE (14% vs. 23%). For sectioned otoliths however, agreement improved substantially when very poor advanced readings (from 2 readers) —showing noticeable drift—were removed, increasing PA to 82%.

Age-class patterns also differed between the two preparation types for the advanced readers. The agreement for sectioned otoliths remained comparatively stable (at a low level) across all age classes (60–70%). In contrast, broken otoliths showed a progressive decline in agreement with age: PA dropped from 84% at age 1 to 50% at age 7, indicating increasing difficulty or inconsistency in interpreting older fish using the broken preparation method.

4. Progress and Recommendations:

This workshop concluded poor overall agreement. The same issues as in 2016 including; interpretation of first annual ring due to split rings and the wide range of growth that can occur, difficulties interpretating the edge and misinterpretation of split rings and Humphries shadow. This highlighted that training of new readers is important as is repeated calibration is necessary among advanced readers. Recommendation for a workshop to cover this in 2026.

Red Mullet [*Mullus barbatus*] Age Reading Exchange [(SmartDots event ID 1914)] and Red Mullet [*Mullus surmuletus*] Age Reading Exchange [(SmartDots event ID 1919)], coordinated by Pierluigi Carbonara (COISPA) and Andrea Massaro (APLYSIA)

1. In 2024, ICES WGBIOP (ICES, 2024. WGBIOP, outputs from 2023 meeting) identified and recommend as necessary to considerate otolith exchange for red mullet (*Mullus barbatus*) and striped red mullet (*Mullus surmuletus*; WKVALMU2). Pierluigi Carbonara (Fondazione COISPA, Italy) and Andrea Massaro (APLYSIA, Italy) were responsible to organising a *Mullus* species (*Mullus barbatus* and *Mullus surmuletus*) otolith exchange.
   Three age reading workshops were conducted in 2009, 2012, and 2017 (ICES, 2009; 2012; 2017), each of which was preceded by an exchange. A total of 22 readers participated in WKVALMU2: 7 Advanced and 15 Basic for *Mullus barbatus*, and 8 Advanced and 14 Basic for *Mullus surmuletus*, representing 8 countries (Albania, Croatia, France, Greece, Italy, Slovenia, Spain and UK).

2. A total of 440 otoliths (258 *M. barbatus* and 182 *M. surmuletus*) were collected from Mediterranean Sea (GSAs: 1, 5, 6 7, 9, 11, 16, 18, 19, 22) and Atlantic (ICES: 7, 8c, 9a). To reduce subjectivity in age reading, length information was not provided. Additionally, to avoid bias associated with different ageing schemes, readings were conducted by marking all annual translucent rings on the otoliths and recording the edge type (Opaque/Translucent). Age assignment will be carried out during the data processing phase by applying a standardized ageing scheme to all readings. Results: For *Mullus barbatus*, the results showed low precision, with percent agreement (PA) of 49.0% for all readers and 59.0% for advanced readers. The coefficient of variation (CV) ranged from 50.0% (all readers) to 45.0% (advanced readers), while the average percent error (APE) decreased from 38.0% (all readers) to 32.0% (advanced readers). 9.1% of the annotations received an AQ3 score. Results for *Mullus surmuletus* show the same low precision with PA 41.0% (all readers) and 44.0% (advanced readers), CV 55.0% (all readers) and 53.0% (advanced readers), APE 49.0% (all readers) and 41.0% (advanced readers). An AQ3 score was assigned to 10.9% of the annotations.

3. Conclusions, recommendations, further actions:
   These results may be explained by the difficulty in identifying the position of the first growth increment and the high incidence of false rings. Moreover, compared to the previous exchange, the marked decrease in precision, in terms of PA, for both species could be related to reader turnover and the extended time interval between exchanges.

   Recommendations from the workshop highlighted a need to standardize protocols to include date of birth, ageing scheme and methodology. Updating the otolith reference collections and associated image databases was also recommended to reflect standardized methods and current best practices. Further discussion with the aim to propose appropriate validation methods for Mullus species can be concluded from this exchange.

North Sea sprat (*Sprattus sprattus*) Age reading Exchange [SmartDots event 1910] Coordinators/organizer Denmark: Julie Davies

- ICES stock name: Sprat (*Sprattus sprattus*) in Division 3.a and Subarea 4 (Skagerrak, Kattegat, and North Sea)
- Link to the SmartDots event 1910 published report: https://smartdots.ices.dk/sampleImages/2025/1910/2025%20NorthSeaSprat%20Full%20Age%20Report%20Event1910.pdf
- Stock code: spr.27.3a4
- Aphia ID: 126425
- Ecoregion: Greater North Sea Ecoregion
- Assessment WG: WKBHERSPR Benchmark workshop on Herring and Sprat stocks
- ICES stock data category: 1
- Assessment type: Age-based analytical assessment (SMS; ICES, 2024b), quarterly time-steps that uses catches in the model
- Last exchange: 2016
- Benchmark planned: 2025
- Background/reasons/aims for event: To address potential issues in age determination – 1st wr, edge type. Same issues as in 2016. Re-reading of 2016 agreed age set showed positive bias. Some inter-lab image sharing and checking in recent years.

- Sample size: 112 (48 samples with 2 images per sample, Reflected light (DTU) and transmitted light (SWE)
  16 "agreed age" otoliths from 2016 WKARSPRAT.
- Method: Whole otoliths
- Number of participants: 17 age readers in total
  - 9 advanced assessment readers 9 (DK, NOR, SWE, NLD, SCOT, DE)
  - 8 basic age readers (FRA, POL)

Results:
PA, CV, APE for advanced and all readers separately (AQ/QS 1 and 2 only).

Age readers	N Sample	CV	PA	APE
Advanced age readers (N=9)	112	28%	77%	18%
All age readers (N=8)	112	35%	71%	25%

Conclusion

PA only 73% at modal age 0 and 2. High positive bias at modal age 0 and 1. SCT and NLD - high bias across all ages. Narrow translucent zone in the innermost opaque zone. Disagreement and uncertainty on how to interpret the growth structures laid down in the first years. Banded appearance of 1st translucent zone. Uncertainty on how/when to count a translucent zone at the edge (June/July). Recommendation from WKBHERSPR for an age reading workshop in 2026.

Actions needed

Training new readers is very important. New readers not familiar with the agreed guidelines. Some inter-lab checks. Repeated calibration is necessary. Presentation at the DECW for WKBHERSPR. Concerns about data quality for assessment. Which readers are reading most for assessment? New model will include sensitivity testing but not before benchmark. New workshop with deeper otolith analysis. Resolution needs to be finalized with ToR's, dates, location and co-chairs.

Blue whiting (*Micromessistius poutassou*) Otolith Workshop [SmartDots event 1922 -1923 Exchanges for the northern and southern areas, WKARBLUE4]

ICES stock name**:** whb.27.1-91214

Assessment WG: Working Group on Widely Distributed Stocks (WGWIDE; ICES 2023a).

Stock category: 1 (ICES, 2023b). Age-based analytical assessment (SAM; ICES, 2021).

Aphia ID: 713132

Background

Blue whiting (*Micromessistius poutassou*) is a widely distributed stock (whb.27.1-91214). The exchange for 2025 took place in Tórshavn 24th June – 27th June in the Faroes Islands. Previous exchanges and workshops were performed in 2013, 2017 and 2021. The species is widely distributed along the Northeast Atlantic with the blue whiting from the northern areas of the stock (North of the British Isles) has a slower growth than in the more southern areas (south of the British Isles). Otoliths from the northern areas (2.a, 5.b, 6.b and 7.c) have proven more difficult to ageread. Other main issues involve interpreting the position of the first annual growth ring, false or split rings and interpretation of the edge. There are also many new age readers. Due to these factors, there is a need to regularly conduct intercalibrations on age reading for stock assessment. New AEM for the next BW benchmark is planned for 2026/2027.

Target PA >70%

Northern areas (SmartDots event: 1922)

The sample collection consisted of 344 otoliths from blue whiting from ICES Div. 2.a, 5.a, 5.b, 6.a, 6.b, 7.c and 7.k. The method for age estimation was whole otoliths soak for 12-24 hours before reading using reflected light. In total, 24 age readers from seven countries participated in the exchange and 14 were advanced age readers.

Results:

Table 1. Results for the otolith Exchange – northern areas (SmartDots event: 1922).

ICES Div.	N samples	CV	PA	APE	Modal age
All Northern	344	26%	67%	16%	0-10
2a	63	27%	71%	16%	0-10
5a	40	17%	74%	11%	0-8
5b	59	37%	59%	25%	0-10
6a	89	18%	72%	12%	0-10
6b	25	14%	70%	8%	3-10
7c	45	13%	79%	8%	3-10
7k	23	13%	75%	10%	2-8

Southern areas (SmartDots event: 1923)

Blue whiting (*Micromessistius poutassou*) is a widely distributed stock (whb.27.1-91214). Previous exchanges and workshops took place in 2013, 2017 and 2021. Due to new age readers there is a need to regularly

conduct intercalibrations on age reading for stock assessment. New AEM for the next BW benchmark (planned for 2026/2027).

The sample collection consisted of 214 otoliths from blue whiting from ICES Div. 8.c and 9.a. The method for age estimation was whole otoliths soak for 12-24 hours before reading using reflected light. In total, five age readers from two countries participated in the exchange and two were advanced age readers.

ICES Div.	Type of age readers	N samples	CV	PA	APE	Modal age
All Southern	All	214	10%	90%	4%	0-7
8c	Advanced	130	1%	99%	0%	0-7
9a	Advanced	83	1%	99%	0%	0-5

Progress from previous years

Progress from previous years. Results from advanced age readers:

WK year	ICES Div.	N samples	CV	PA	APE	Modal age
2021	All	407	19%	69%	11%	0-13
2021	N	190	20%	69%	11%	0-13
2025	N	344	25%	71%	15%	0-10
2021	S	217	24%	79%	14%	0-7
2025	S	214	1%	99%	0%	0-7

Conclusions

WKARBLUE4 recommends conduct regular age reading intercalibration exchanges.

Further actions needed
WKARBLUE5 planned for 2029.

Coordinators
Patrícia Gonçalves (IPMA, Portugal) and Jane Godiksen (IMR, Norway)

References:
ICES. 2021. Working Group on Widely Distributed Stocks (WGWIDE). ICES Scientific Reports. 3:95. 874 pp. http://doi.org/10.17895/ices.pub.8298
ICES. 2023a. Working Group on Widely Distributed Stocks (WGWIDE). ICES Scientific Reports. 5:82. https://doi.org/10.17895/ices.pub.24025482
ICES. 2023b. Advice on fishing opportunities. In Report of the ICES Advisory Committee, 2023. ICES Advice 2023, section 1.1.1. https://doi.org/10.17895/ices.advice.22240624

Horse Mackerel, Mediterranean Horse Mackerel and Blue Jack Mackerel. [*T. trachurus, T. mediterraneus and T. picturatus*] [Otolith exchange, previous to WKARHOM5] [(WKEvent ID: 2926 - 2927 - 2928)]

In 2022, the participants in WKARHOM 4 agreed to convene a new full otolith exchange for horse mackerel, Mediterranean horse mackerel, and blue jack mackerel in 2025, prior to the potential celebration of WKARHOM5 (Workshop on Age Reading of *Trachurus trachurus, Trachurus mediterraneus, Trachurus picturatus*). This need was approved during by the WGBIOP (ICES, 2024).

Andrea Massaro (APLYSIA, Italy), Alba Jurado-Ruzafa (IEO-CSIC, Spain) and Justine Diaz (IMR, Norway) have coordinated the age reading exchange for *Trachurus* species *(T. trachurus, T. mediterraneus and T. picturatus)*. Results from previous workshops and exchanges (Bolle *et al*., 2011; ICES, 1999, 2012, 2015, 2018, 2023; Mahè *et al*., 2015) consistently indicated low precision in age readings, with particularly significant discrepancies related to the otolith preparation methods used. A total of 22 readers from 8 countries (Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal and Spain) have participated in the previous otolith exchange to WKARHOM5: 11 Advanced and 9 Basic for *Trachurus trachurus*, 1 Advanced and 14 Basic for *Trachurus mediterraneus* and 16 basic for *Trachurus picturatus*. To reduce subjectivity in age reading, length information was not provided. Additionally, to avoid bias associated with different ageing schemes, readings were conducted by marking all annual translucent rings on the otoliths and recording the edge type (Opaque/Translucent).

Globally, these results may be explained by the difficulty in identifying the position of the first growth increment and the high presence of false rings, especially in *T. mediterraneus* and *T. picturatus*. Likewise, the "stacking" phenomenon of growth zones towards the otolith margin increases the difficulty in reading of old specimens. Moreover, compared to the previous WK, a decrease in precision, in terms of PA, mainly in *T. mediterraneus* and *T. picturatus,* could be related to the considerable number of participants with low or any experience, reflecting the increasing interest on training new readers and providing interspecies input for age determination.

Trachurus trachurus

A total of 221 otoliths' images (167 whole and 54 sliced) were provided by participants including the Mediterranean Sea (GSAs: 8, 11, 20) and the East Atlantic Ocean (ICES 27.4.a, 27.6.a, 27.8.c, 27.9.a). Noticeable differences were obtained comparing otoliths' preparation, with a coefficient of variation (CV) markedly lower in sliced otoliths (21%). However, the percent agreement (PA) remained similar between preparations in the readings performed by advanced readers. The percent agreement (PA) for whole otoliths was 54% and 56% for basic and advanced readers, respectively. For sliced otoliths, the PA was 48% for basic readers and 55% for advanced readers. A total of 3.2% of the annotations received an AQ3 score.

Trachurus mediterraneus

A total of 103 otoliths' images were provided by participants from the Mediterranean Sea (GSAs: 1, 6, 11, 20). The results of the readings showed low precision, with percent agreement (PA) of 55.0% for all readers, a coefficient of variation (CV) of 42.0% (all readers), and an average percent error (APE) of 31.0%. An AQ3 score was assigned to 5.4% of the annotations.

Trachurus picturatus

A total of 81 otoliths' images were provided by participants from the Ligurian and North Tyrrhenian Sea in the Mediterranean Sea (GSA 9) and from Azores grounds in the East Atlantic Ocean (ICES 27.10.a.2). The results of the readings showed showed low precision, with percent agreement (PA) of 50.0% for all readers (all of them with basic expertise), a coefficient of variation (CV) of 39.0% (all readers), and an average percent error (APE) of 29.0%. A total of 14.4% of the annotations received an AQ3 score.

Gulf of Bothnia herring her.27.3031 (*Clupea harengus*) Age reading Exchange [SmartDots event 1849]. Coordinators/organizers: Sweden: Yvette Heimbrand and Annelie Hilvarsson SLU Aqua, Finland: Annie Pursiainen and Sami Vesala LUKE. SmartDots

- ICES stock name: Herring (*Clupea harengus*) in subdivisions 30 and 31 (Gulf of Bothnia)
- Stock code: her.27.3031
- Aphia ID: 126417
- Ecoregion: Baltic Sea Ecoregion
- Assessment WG: Baltic Fisheries Assessment Working Group (WGBFAS)
- ICES stock data category: 1 (ICES, 2023)
- Assessment type: Age-based analytical assessment, Stock Synthesis (ICES, 2024a)
- Background: Sweden and Finland age herring from the same area (Bothnian Sea, SD 30 and 31) with the same method. Regular age calibrations are recommended for quality assurance for the age based analytical assessment. Previous results (2019) for 100 otolith images overall good with >80% PA but with high CV for age class 0 and 1 due to bad quality images for age class 0 and 1. Variability of estimated weight-at age has affected the estimation of spawning-stock biomass (SSB) in recent years that could be driven by the availability of important zooplankton prey (ICES. 2025). However, to assure age data quality, regular age calibrations are important.

- Sample size: 200 otolith images
- Method: Sectioned and stained with toluidine blue
- Number of participants: 12 age readers in total
    - 6 advanced assessment readers (4 from Sweden SLU Aqua and 2 from Finland LUKE)
    - 6 basic readers (2 from Finland LUKE and 4 from The University of Turku)

Results:

PA, CV, APE for advanced and all readers separately (AQ/QS 1 and 2 only).

Age readers	N Sample	CV	PA	APE
Advanced age readers (N=6)	200	10%	91%	4%
All age readers (N=12)	200	14%	87%	7%

- Otoliths with AQ3 score = 0,21%
- Agreement threshold PA > 80% considered as consensus
- Conclusion: The results from the age calibration show good agreement among age readers from both Sweden and Finland.
- Actions needed: Organize a meeting with participating age readers to discuss results and select otolith images with 100% PA to include in a new training reference collection for training of new readers and self-testing of experienced readers. Continue with regular age calibrations.
- Coordinators/organizers: Sweden: Yvette Heimbrand and Annelie Hilvarsson SLU Aqua, Finland: Annie Pursiainen and Sami Vesala LUKE. SmartDots

References:

ICES. 2023. Advice on fishing opportunities. In Report of the ICES Advisory Committee, 2023. ICES Advice 2023, section 1.1.1. https://doi.org/10.17895/ices.advice.22240624

ICES. 2024a. Baltic Fisheries Assessment Working Group (WGBFAS). ICES Scientific Reports. 6:53. 670 pp. https://doi.org/10.17895/ices.pub.25764978

ICES. 2025. Herring (*Clupea harengus*) in subdivisions 30 and 31 (Gulf of Bothnia). In Report of the ICES Advisory Committee, 2025. ICES Advice 2025, her.27.3031, https://doi.org/10.17895/ices.advice.27202623

Lemon sole (*Microstomus kitt*) in ICES divisions 27.4c and 27.7d Age Reading Exchange (SmartDots event ID 115) coordinated by Joanne Smith (Cefas, UK)

This lemon sole age reading exchange took place between July to September 2019 using SmartDots software. No previous age validation has been done on lemon sole from ICES divisions 27.4 and 27.7d and they can be a difficult species to age. Therefore, it was advised by readers and Working Group on Biological Parameters (WGBIOP 2018) to have a full-scale otolith exchange in order to identify and resolve age interpretation differences between readers, laboratories and methods. A total of nine participants, from six countries were involved in the exchange. A decision was made to use 27.4 and 27. 7d as this is an age assessed stock and including additional areas with so many methods would result in a large number of sets. Following WGBIOP Guideline for Otolith Exchanges (2018) a set of 108 whole, 64 sectioned and 64 broken and burnt otoliths were selected and uploaded on to the SmartDots application. The samples were provided from UK (Cefas) and Belgium (ILVO).

The objectives of the exchange were: -

- Evaluate the accuracy and precision in age reading of lemon sole 4 and 7d. –
- Identify any issue in reading lemon sole 4 and 7d.
- Identify the accuracy of reading lemon sole IV and VIId using different preparation techniques.

Results

The statistics representing age reading performance were calculated for all readers combined and for experienced readers only. As expected, agreement was higher, and variance (APE and CV) was lower for advanced readers compared to all readers regardless of preparation method.

All areas were included, and calculations were carried separately for each preparation methods – section, whole and broken/burnt. Due to a small number of readers per preparation method and to allow us to compare accuracy of each preparation method, readers were asked to read all three methods – but it should be noted that not all readers use all three methods to age this species.

In all cases the statistics were significantly better for sectioned otoliths than for whole or broken/burnt. The average percentage agreement of 62% and CV of 17% was reached by all readers annotating sectioned otoliths. There was a large improvement when only advanced readers were combined: PA=90%; CV=3%, however there were only two readers who were from the same institute.

The average percentage agreement of 49% and CV of 25% was reached by all readers annotating whole otoliths. There was a slight improvement when only advanced readers were combined: PA=51%; CV=25%, The average percentage agreement of 25% and CV of 29% was reached by all readers annotating burnt otoliths. It should be noted that none of the readers were expert in this method and poor images made annotation of ages more difficult. The Uk is the only institute that reads lemon sole using this method and

future exchanges should use otoliths rather than images, until photography of burnt structures can be improved.

Usually, a CV of 5% is set as a threshold for sufficient data quality (Campana 2001). The results of present exchange indicate the proposed threshold statistics are not achieved by all experienced readers. This is related to the interpretation differences and the main discrepancies were caused by uncertainties in the first ring, particularly in whole otoliths. However, it was pleasing to see that readers seemed confident in distinguishing between true and shadow rings.

Institutes tend to use different methods of reading, Denmark and Iceland read whole, Belgium reads whole/sectioned Netherlands reads sectioned and UK reads whole/broken/burnt. Most readers read all three methods to help us get and understanding of accuracy between methods. Due to the number of different of methods used it was decided to include all readers in the results, this could have introduced bias to the results but with the limited number of readers per method it was decided this was the best approach. Trainee readers were also left in as removing them did not seem to have a positive effect on the outcome.

Discussion/Conclusion

Institutes tend to use different methods of reading, Denmark and Iceland read whole, Belgium reads whole/sectioned Netherlands reads sectioned and UK reads broken/burnt.

Percentage agreement between all readers of whole lemons was 49%, sectioned 62% and burnt 25%.

Removing readers R08, R16 and R26 (low percentage agreement) from the sectioned results increased percentage agreement to 85%.

Removing readers R12 and R14 (lower percentage agreement) from the whole results increased the agreement to 62%, which is still very low.

Percentage agreement was very low for burnt otoliths, but this may be due to poor images of the burnt otolith. It is very difficult to take a clear image with water on the surface, but this problem is being looked into to address better ways to photograph such structures.

Percentage agreement of expert readers for section was very good at 90%, however this included just two readers who were from the same institute. Percentage agreement for expert readers for whole otoliths was much lower at 51%.

The results from this exercise show sectioned method provides a higher quality preparation for age determination and that results obtained from reading the sectioned otoliths are more reliable than those obtained by the two other methods (whole/ broken and burnt).

The results also showed that using the preparation method of whole otoliths, can lead to an underestimation of age readings by some readers, with particular issues in interpreting the first ring. There cannot be any comparison of results to previous exchanges as this is the first carried out for lemon sole.

The low percentage agreement between readers and institute highlights the importance of this exchange. During this exchange many readers encountered the same issue with some including and some excluding the first ring. Comparisons of the same otolith prepared in different ways showed that the age of whole otoliths was often underestimated as compared to sectioned or broken burnt otoliths. This must be noted however that this was a small sample size and a greater number would be needed to make a true comparison. Following on from this exchange it would be recommended for a workshop to be carried out to discuss discrepancies in age readings, particularly with the first ring and possibly an otolith chemistry study

Sardine (*Sardina pilchardus*) Otolith Exchange (SmartDots event ID 2950), co-ordinated by Cristina Bultó (CSIC) and Andreia Silva (IPMA)

This otolith exchange was carried out for ICES areas 27.8.c, 27.9.a and in GFSA areas GSA01, GSA03, GSA04, GSA06, GSA8, GSA09 and GSA19. The exchange was a request from the WGHANSA. The assessment uses age compositions from subdivisions 27.8abd and 27.8c9a, since last intercalibration was in 2017 readers have changed significatively until our days, several experienced readers have retired, new readers have joined multiple laboratories. Also, a new sardine stock emerged in the Celtic Seas and English Channel (Subarea 7) in 2017, and no samples from this stock were available in the last intercalibration. WGHANSA recommended a new intercalibration and discussion workshop on age reading of European sardine (Sardina pilchardus) in the NE Atlantic.

The last sardine ageing workshop was held in 2019 (Lisbon), where Percent Agreement (PA) reached 75.6% for Atlantic areas and 69.7% for the Mediterranean. The background and aim was to check the level of agreement between age readers and be able to identify possible issues. A total of 20 readers participated (3 advanced from Atlantic areas and 5 advanced from Mediterranean areas), representing countries (7 European and 2 African). Two advanced readers from Atlantic areas (delivering data for stock assessment) provided photos but did not submit readings and there are no advanced readers participated from area 27.8.b. The exchange was conducted through the SmartDots platform, and a preliminary analysis was performed using R scripts (Silva and Soares, 2022). The list of participants can be found below.

Participants list and Expertise:

Country	Expertise
UNITED KINGDOM	Basic
SPAIN	Advanced
ALGERIA	Basic
MOROCCO	Basic
UNITED KINGDOM	Advanced
ALGERIA	Basic
GREECE	Advanced
GERMANY	Basic
SPAIN	Basic
ITALY	Advanced
FRANCE	Basic
ITALY	Basic
PORTUGAL	Advanced
ITALY	Advanced
SPAIN	Advanced
PORTUGAL	Advanced
PORTUGAL	Basic
SPAIN	Basic
SPAIN	Basic
SPAIN	Advanced

Results:

The otolith exchange exercise was successfully completed, with consistent participation and results across both Atlantic and Mediterranean regions. Preliminary analyses showed acceptable levels of agreement (PA) and variability (CV), providing a clear overview of differences in reader consistency between regions (Atlantic PA: 73.1%; CV: 20.7%; Mediterranean PA: 71.1%, CV 35.7%).

Conclusions and recommendations:

This workshop concluded new stock areas require calibration with recently joined readers, and further studies on growth are needed. Consistency across laboratories should continue to be monitored, especially with the integration of new participants. Recommendations include ensuring advanced readers are represented in all areas (particularly Atlantic subareas 27.9.a and 27.8.c). This workshop highlighted the need to improve the balance of advanced reader coverage between Atlantic and Mediterranean regions. There is a need to secure support from a second coordinator in future to share responsibilities and ensure timely delivery. A plan to follow up this workshop to consolidate results and maintain intercalibration across laboratories is proposed.

2024 Atlantic Mackerel (*Scomber scombrus*) 127023 [Histology screening and fecundity Exchange] [WKAEPM2]

Coordinated by Ewout Bloom (Wageningen University and Research, Netherlands) and Hannah Hola (Marine Scotland Science, UK).

A calibration exchange among the participating institutes on western component of Atlantic mackerel stock (mac.27.nea) ovary screening, fecundity, atresia, and postovulatory follicle (POF) analysis was carried out prior to the WGMEGS survey.

These measurements are used to estimate the SSB index from the survey, which are then used in the assessment conducted by WGWIDE (Working Group on Widely Distributed Stocks). The assessment uses SAM, an age-based analytical model, for both the western and North Sea components of Atlantic mackerel (subareas 1-8 and 14 and in divisions 9.a, 12.a, and 12.b - Northeast Atlantic and adjacent waters).

A benchmark for Northeast Atlantic mackerel was held in the first quarter of 2025. Among the list of priority issues for the benchmark, four were related to MEGS and needed to be addressed: 1) alternative ways of estimating potential fecundity and atresia; 2) Calculate the variance for AEPM fecundity; 3) Provide SSB using the DEPM; 4) integrate the results of the North Sea DEPM survey with the western and southern AEPM surveys.

The aim was to find out about possible differences between institutes so that uncertainty is known and improvements could be proposed and updated in the survey manual SISP-5 (ICES, 2019). At the same time, the benchmark request, particularly the first one, was also considered by the workshop.

17 participants (including, both advanced and other readers jointly) from 6 countries took part in the exercises, which involved the analysis of 27, 2, 3, and 25 samples for screening, fecundity, atresia, and POF staging, respectively. To statistically evaluate the results, the Average Percent Error (APE) and the Coefficient of Variation (CV) were used as precision indices. Percent Agreement (PA) was also employed to facilitate comparison with previous calibration exercises.

Histological screening: the percentages of agreement were above 50% in all cases. For the previtellogenic, early vitellogenic, late vitellogenic and migratory nucleus oocyte stage the percentages of agreements were

above 87,5% in between 78% and 100% of the samples. For the presence of egg stages there were agreements above 87, 5% in 89% of the samples while massive atresia had agreements above 87.5% for 93% of the cases. The lowest agreements were obtained for the presences of early alpha atresia, postovulatory follicles, spent ovaries and hydration states (52%, 67%, 70% And 74% respectively).

POF staging: The results showed that the readers most frequently reached an agreement level of 50%, which is considered low. In most samples, the discrepancies between readers were limited to adjacent POF stages. Notably, when these results were translated from POF Stages to daily cohort stages, the most common agreement level increased to 75%. The mean coefficient of unalikeability was approximately 0.55, indicating that the scores were roughly halfway toward a uniform distribution and thus exhibited intermediate variance. This variance minimized (0.29) when scores converged to daily cohort categories.

Fecundity: Using thresholds of 185 μm and 230 μm, the fecundity counts for the prespawning fish ranged from 350 to 442 and from 330 to 415, respectively. These ranges corresponded to a coefficient of variation (CV) of 8.1% for 185 μm and 9.2% for 230 μm.

Atresia: Discrepancy of point counting was rather similar in all the samples. Among the readers there was one reader whose results were more frequently different from the ones from other readers. The evaluation of the results showed that this was specifically due to this person having a different interpretation of the counting rules than the others.

The enhanced results on screening and POF staging compared to previous calibration exercises were due to several key factors: active participation in workshops, accumulated experience from years of sample analysis, the introduction of clearer guidelines and criteria in the manual, and the consistent use of standardized, calibrated measurement tools. Additionally, the use of digital platforms for collaborative discussions helped resolve doubts, align interpretations, and foster a shared commitment to accuracy. To improve consistency in fecundity measurements, participants reached agreements on how to handle oocytes that are not perfectly circular, broken, or partially covered. Differences in oocyte counts between the 185 μm and 230 μm thresholds across institutes will be further investigated using artificial beads for calibration. The atresia manual was revised following discrepancies observed during the atresia exercise.

ICES. 2019. SISP 6 - Manual for mackerel and horse mackerel egg surveys, sampling at sea. Version 2.2. Series of ICES Survey Protocols. 82 pp. http://doi.org/10.17895/ices.pub.5140

2024 North East Atlantic Mackerel (*Scomber scombrus*) and Horse mackerel (*Trachurus trachurus*) [(WKMACHIS2 and SmartDots ID 1883 and 1885)]

General information

The Ices Workshop on Mackerel, Horse mackerel and Hake Eggs Identification and Staging (WKMACHIS2) was held from 21-25 October 2024 at TI-SF, Bremerhaven, Germany. It dealt with mackerel (127023; mac.27.nea) and horse mackerel (126822; hom.27.2a3a4a5b6a7a-ce-k8, hom.27.4bc7d, hom.27.9a) for the surveys I4189 and I1582. The data are important input for the ICES expert groups Working group on Mackerel and Horse mackerel Egg surveys (WGMEGS) and Working Groups for Widely Distributed Species (WGWIDE). In 2025 there was a benchmark for mackerel (WKBMACNSSH) and in 2024 there was a benchmark for horse mackerel (WKBHMB). This workshop was held for plankton analysts involved in the 2025 ICES coordinated mackerel and horse mackerel egg survey (MEGS), with the aim of standardizing and calibrating the identification and staging of eggs from the target species mackerel, horse mackerel, hake and ling.

Method.
Participants analysed the same 513 eggs across two SmartDots events (one pre-workshop) and one physical event. Results were collated, processed and quality checked, and entered into standard Excel evaluation sheets. These sheets were utilized to identify species, stages, and individual eggs with low agreement scores. These eggs were presented on screen to demonstrate key diagnostic indicators and identify causes of disagreement and misidentification. In this way, the egg staging criteria were reviewed based on shared expertise and group contributions. The workshop was chaired by Ewout Blom (the Netherlands) and Hannah Holah (UK-Scotland). In total 29 persons (4 online) representing 11 institutes from 10 countries participated.

Results.
In the first Smartdots event experts agreed for 94% on stage 1 mackerel eggs, during the second SmartDots event overall agreement for stage 1 mackerel eggs was 98% and for the physical identification exercise agreement was 95% for stage 1 mackerel eggs.

In the first SmartDots event all participats agreed for 93% on stage 1 mackerel eggs, during the second SmartDots event overall agreement for stage 1 mackerel eggs was 99% and for the physical identification exercise agreement was 95% for staging mackerel eggs.

In the first Smartdots event experts agreed for 96% on stage 1 horse mackerel eggs, during the second SmartDots event overall agreement for stage 1 horse mackerel eggs was 96% and for the physical identification exercise agreement was 94% for stage 1 horse mackerel eggs.

In the first SmartDots event alle participants agreed for 95% on stage 1 horse mackerel eggs, during the second SmartDots even overall agreement for stage 1 horse mackerel eggs was 96% and for the physical identification exercise agreement was 92% for stage 1 horse mackerel eggs.

The full report of the workshop:

ICES. 2025. Workshop on Mackerel, Horse Mackerel and Hake Eggs Identification and Staging 2 (WKMACHIS2; outputs from 2024 meeting). ICES Scientific Reports. 7:35. 54 pp. https://doi.org/10.17895/ices.pub.28638974

Conclusions.

- For the first time both the SmartDots application and a physical event were used
- SmartDots worked well but issues occurred when downloading data, i.e. if participants used commas in the comments
- Besides mandatory species anchovy, sardine, Mueller's pearlside and boarfish will also be counted during new surveys.
- Overall agreement for stage 1 eggs of mackerel and horse mackerel was high.