

INTERNATIONAL PACIFIC



HALIBUT COMMISSION

Using artificial intelligence (AI) for supplementing Pacific halibut age determination from collected otoliths

Agenda item 4.2.6:
IPHC-2026-SRB028-INF01
(B. Hutniczak)



Purpose

- To summarize the current knowledge on the use of artificial intelligence (AI) for determining the age of fish from images of collected otoliths
- To provide an update on the ongoing exploratory work to develop an AI-based age determination model for Pacific halibut

The primary objective is to assess the viability of an AI-based approach as a supplement to existing Pacific halibut ageing protocols, while also identifying the remaining steps and requirements necessary for potential operational implementation.



Key progress updates

- Evaluating multiple components of the **deep learning workflow** for Pacific halibut age estimation, including model architectures, image preprocessing, weighting strategies, and training procedures.
- Evaluating the contribution of **auxiliary biological and environmental covariates**, through comparison of image-only and multi-input deep learning models integrating both visual and tabular data.
- **Evaluating model spatial generalization** by comparing age predictions from a model trained on images from one area to those from a different area.
- **[SRB request] Evaluating model temporal generalization** by comparing age predictions from a model trained on images from one year to those from a different year.
- Utilizing **confidence intervals derived from deep ensemble techniques** to assess the model's capability in identifying ambiguous or noisy samples.
- Demonstrating improvements in model performance associated with **expansion of the image database** from 2,799 (SRB026) to 11,107 otolith images.



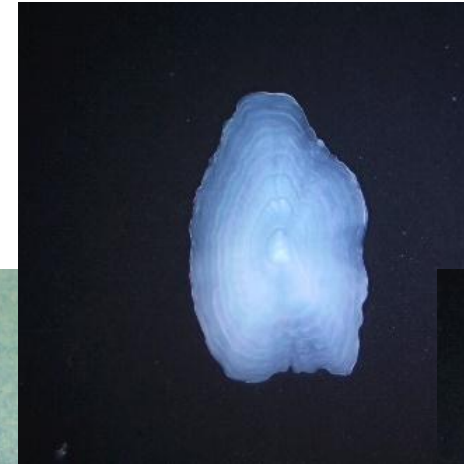
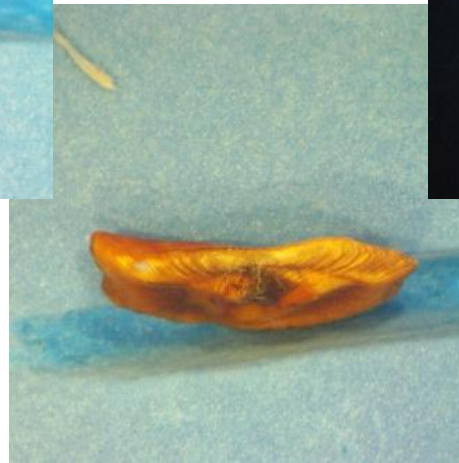
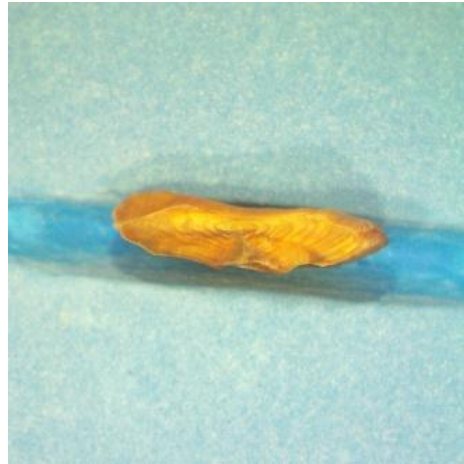
Modeling approach

- Application of **convolutional neural network (CNN) model**, a type of deep learning approach
 - In CNNs, the layers are structured as stacks of filters, each recognizing increasingly abstract features in the data.
- Application of **image regression** - predicting a continuous variable from an image
 - Pacific halibut is a long-living species - oldest Pacific halibut on record were aged at 55 years
- Implementation - TensorFlow and Keras libraries, repurposing **Inception V3** model:
 - Input → InceptionV3 (feature extractor) → Regressor → Output
 - Inception V3 provides more reliable and robust predictions
 - Alternatives: EfficientNet variants (EfficientNetB4, EfficientNetB5, EfficientNetV2 S/M/L) and ConvNeXt; PyTorch framework



Database

- Since 1925, over 1.6 million otoliths have been aged and stored for potential future use. ← unique resource for AI training
- Aged otoliths are sectioned (broken in half) and baked to enhance the contrast between the growth rings, then photographed.
- In addition, IPHC is in the process of creating complementary database comprising labelled images of otoliths captured prior to processing.



- Pictures are taken with AmScope 8.5MP eyepiece cameras



Latest runs results (InceptionV3)

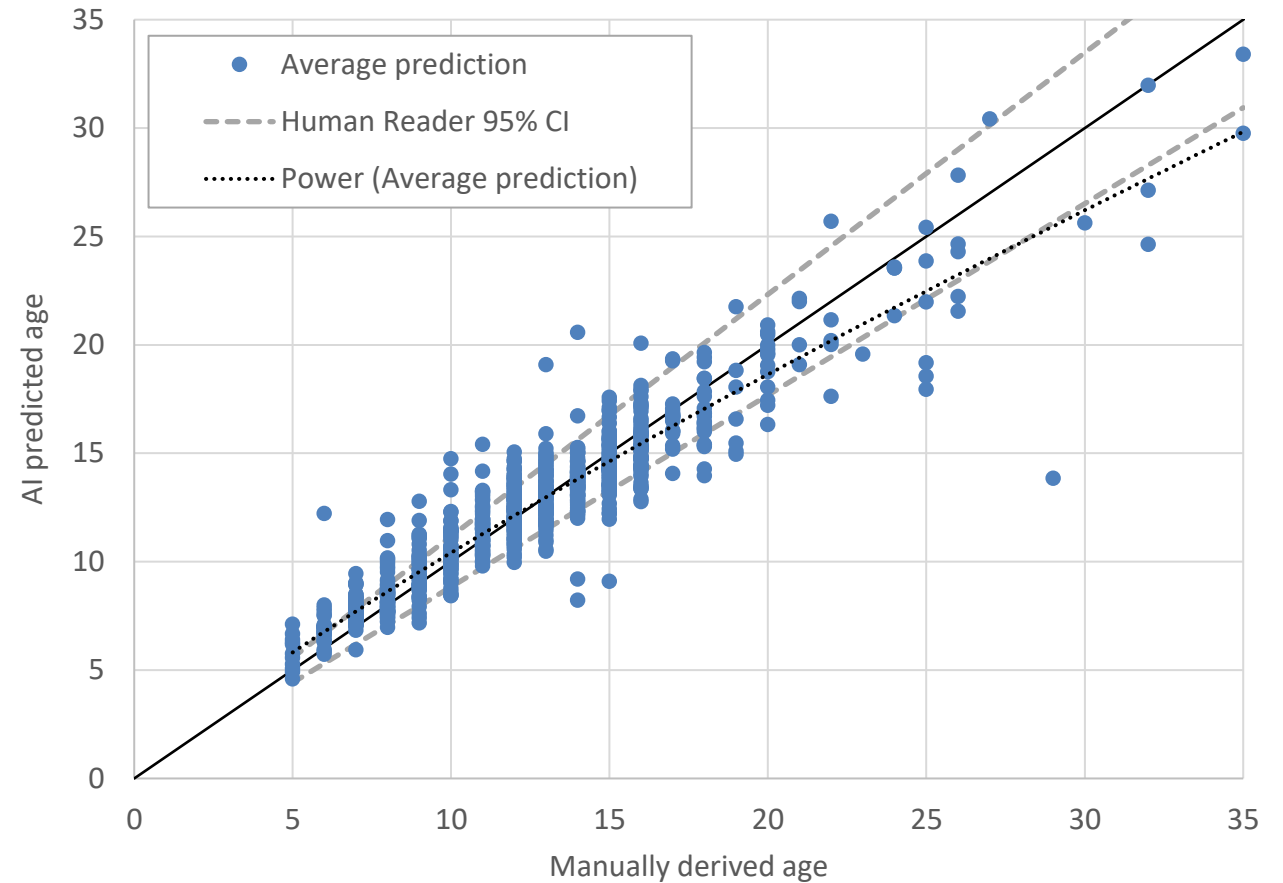
- Main setup (2019 FISS): 11,107 images (7,740 train, 1,367 validation, 2,000 test)
 - Test set is used to assess the performance of the model after training, providing an unbiased evaluation of its generalization capability to new, unseen data.
- Testing the temporal generalization (2024 FISS): 2,931 images
 - All images resized to 400x400 pixels; broken otoliths excluded
 - Epochs max: 600; patience: 60
 - Learning rate: initial at 0.0002 with reduction triggered if validation loss plateaued for 30 epochs
 - Batch size: 8
 - Deep ensemble approach: the model was trained 5 times, each with a different random seed
 - The model trained for an average of 183 epochs. It achieved a normalized MSE of 0.0011 on the validation set



Base run results (1/3)

The model achieved average RMSE in the test set of **1.68** when calculated for rounded results. On average, the ensemble **correctly predicted age for 35.5% individuals**, with an **additional 41.0% being within 1 year of error.**

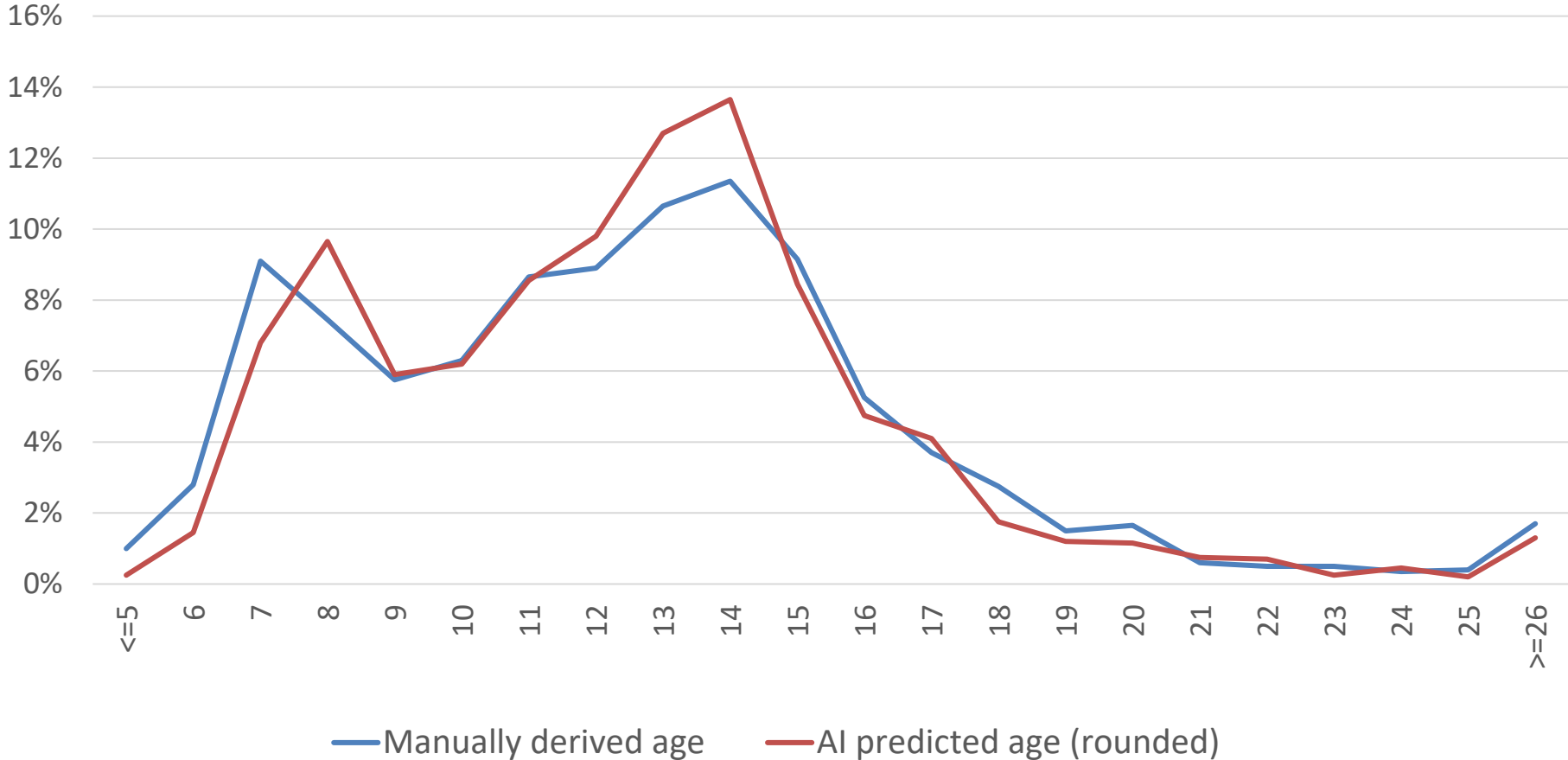
- **Total agreement within 1 year for 76.5% of cases**



Comparison between manually derived age with AI predicted age.



Base run results (2/3)

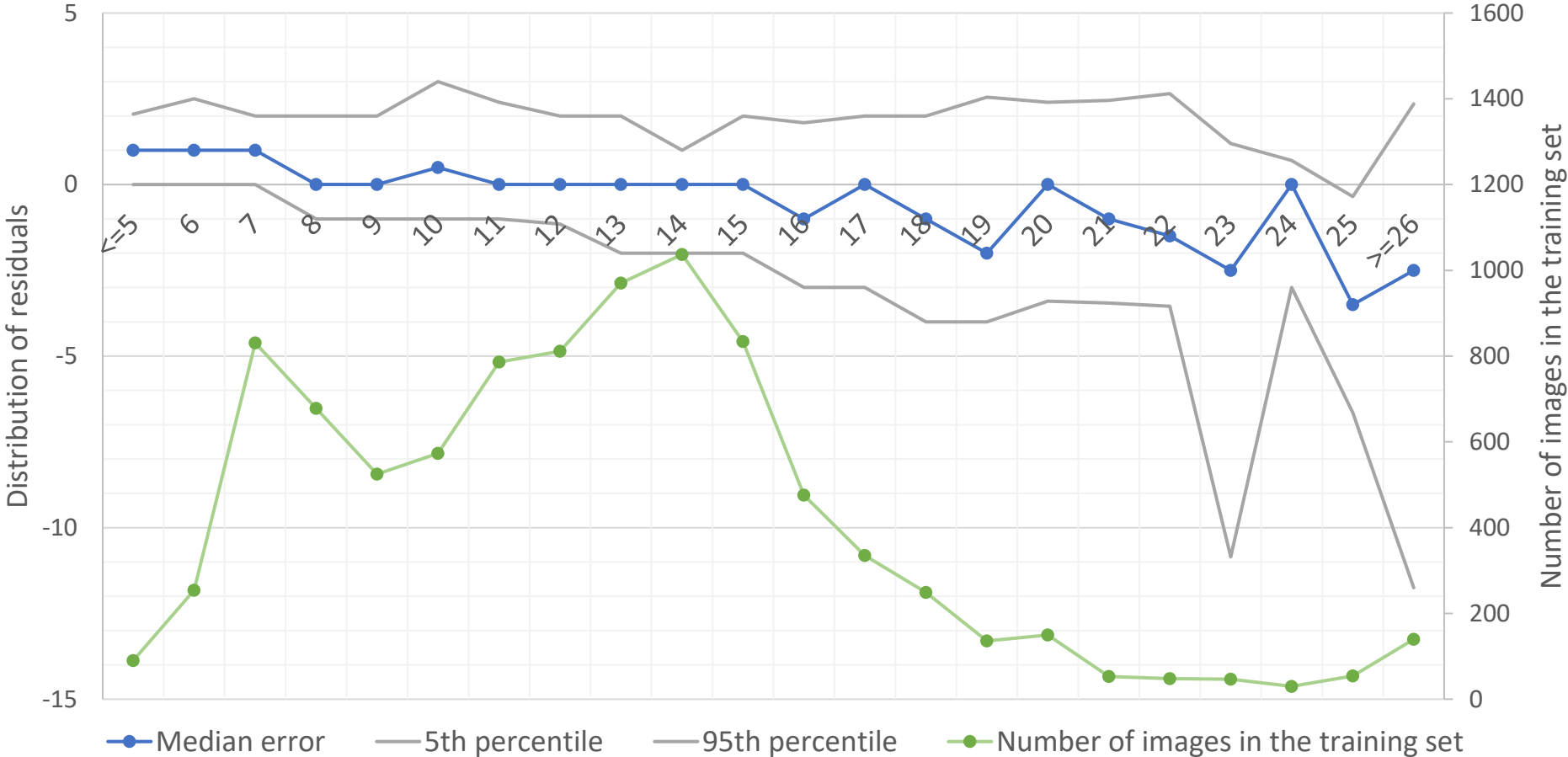


Comparison between manually derived age with AI predicted age – age composition.



Base run results (3/3)

The model tends to underestimate the ages of older Pacific halibut on average, but the statistically significant bias previously observed for age categories 21+ (IPHC-2025-SRB026-10) is no longer apparent.



Distribution on residuals and number of images by age in the test set.

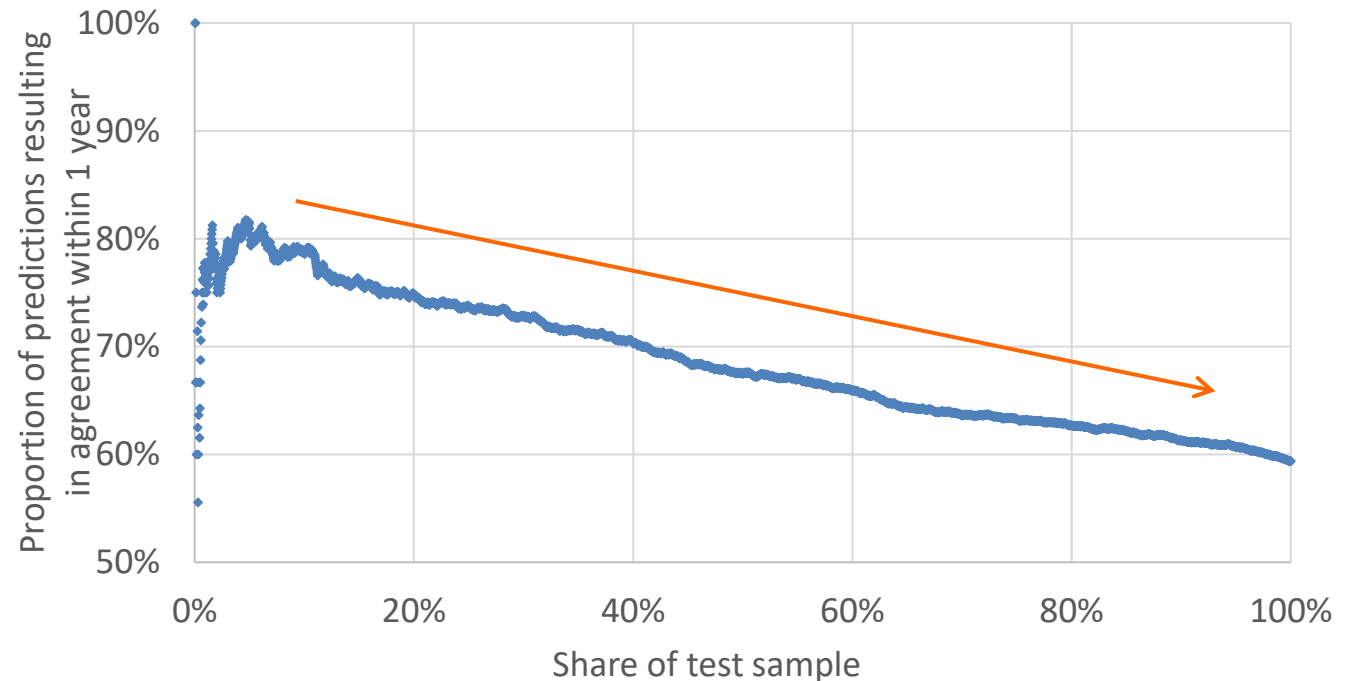


Temporal generalization

- A model trained on 2019 FISS otolith images was evaluated on 2024 data to assess temporal generalization.
- RMSE increased to 2.27 (35% increase compared to within-year predictions).
- The proportion of predictions within ± 1 year of the manually assigned age decreased by 16.9 percentage points.

Fine tuning applied:

- 20% randomly selected images – RMSE = 2.23
- 20% images with the highest standard deviation across ensemble predictions – RMSE = 1.97



Proportion of ensemble predictions within ± 1 year of manual age as a function of cumulative share of the test sample, ordered by prediction uncertainty (standard deviation).



Spatial generalization

- Training on 2,000 randomly selected images including focus area (model 1) and excluding that area (model 2)



Robust performance across IPHC Regulatory Areas - spatial exclusion resulted in limited and inconsistent performance degradation.

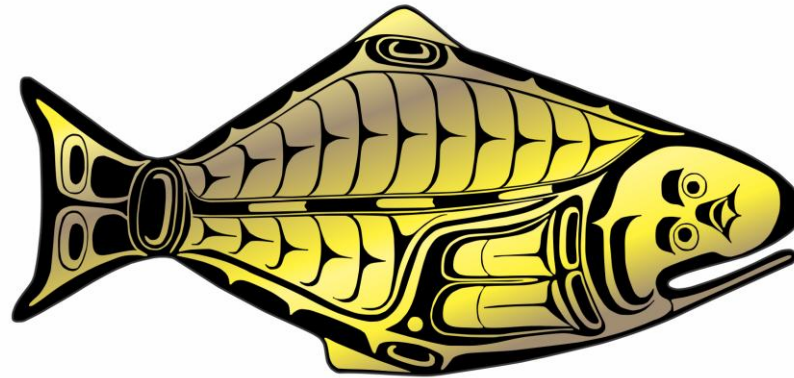


Conclusions

- The ongoing advancement of AI technologies in the field of marine science offers considerable **potential to improve the efficiency of age determination** of Pacific halibut using otolith images.
- Preliminary results presented here suggest that AI could serve as a **promising alternative to the current ageing protocol**, which relies entirely on manual age reading.
 - Continued expansion of the image database improves model performance
 - **NEW:** Use of auxiliary biological and environmental covariates provided modest gains
- **Practical value of the deep ensemble framework** - per-sample uncertainty estimates help identify potentially unreliable predictions, allowing for targeted expert review.
- **Targeted fine-tuning** offers an effective strategy for adapting models to new data, including better temporal generalization.
- An **adaptive approach** will continue to partially depend on trained readers for **capturing temporal changes**.
- **NEW: Spatial exclusion** caused limited and inconsistent performance degradation, suggesting broadly transferable otolith features across IPHC Regulatory Areas.
- **AI is evolving rapidly**, and adapting to new developments may further improve results over time.



INTERNATIONAL PACIFIC



HALIBUT COMMISSION

<https://www.iphc.int/>

