INTERNATIONAL PACIFIC HALIBUT COMMISSION

Supplementing Pacific halibut age determination with deep learning: covariate integration and uncertainty-guided fine-tuning

Barbara Hutniczak

PICES 2026 – FIS

28 Oct 2026



How old is this fish?



Pacific halibut otolith, broken and baked (2019 setline survey; image 2019NCR0330)

Hands up:

A) under 10 years

B) 10 to 20 years

C) over 20 years

How old is this fish?



Pacific halibut otolith, broken and baked (2019 setline survey; image 2019NCR0330)

Hands up:

A) under 10 years

B) 10 to 20 years

C) over 20 years

Expert readers: 32 years | AI ensemble: 32 years

Seven models ranged 29 to 34 years (mean 32.2); the model never saw this otolith in training

Purpose

- Manual fish age determination from otoliths is accurate but time consuming and costly
- **Question: can AI supplement this process?**

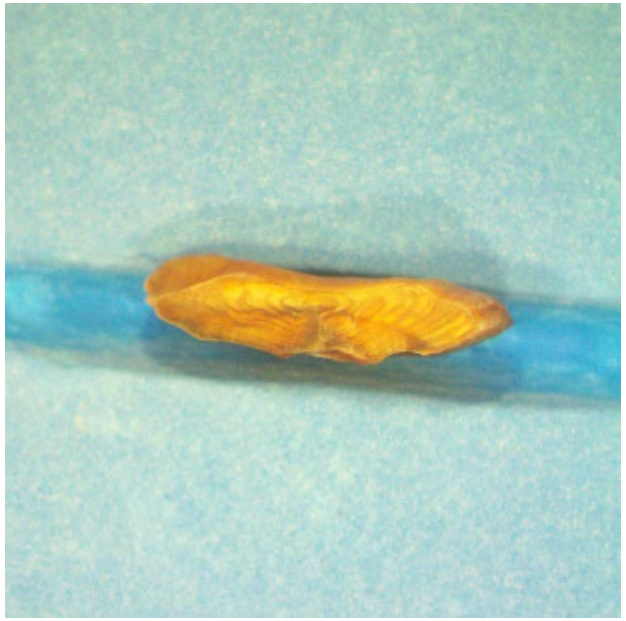
Primary objective: assess the viability of an AI-based approach as a supplement to existing Pacific halibut ageing protocol, and identify the remaining steps and requirements necessary for potential operational implementation.

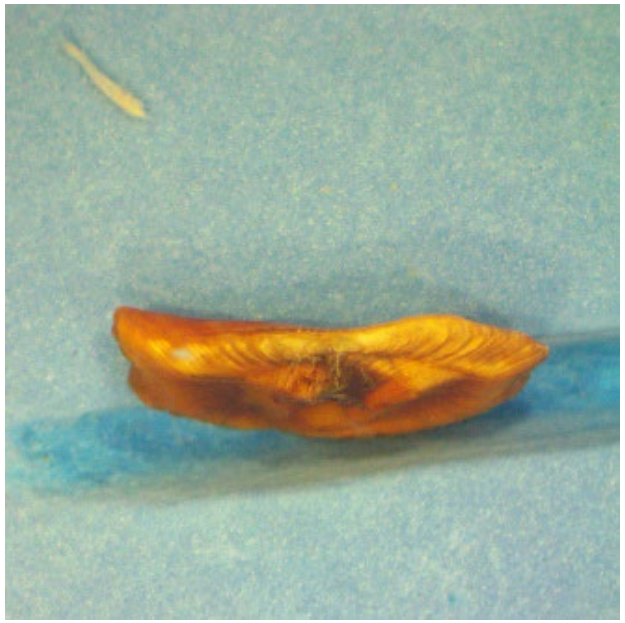
IPHC archives

- 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.



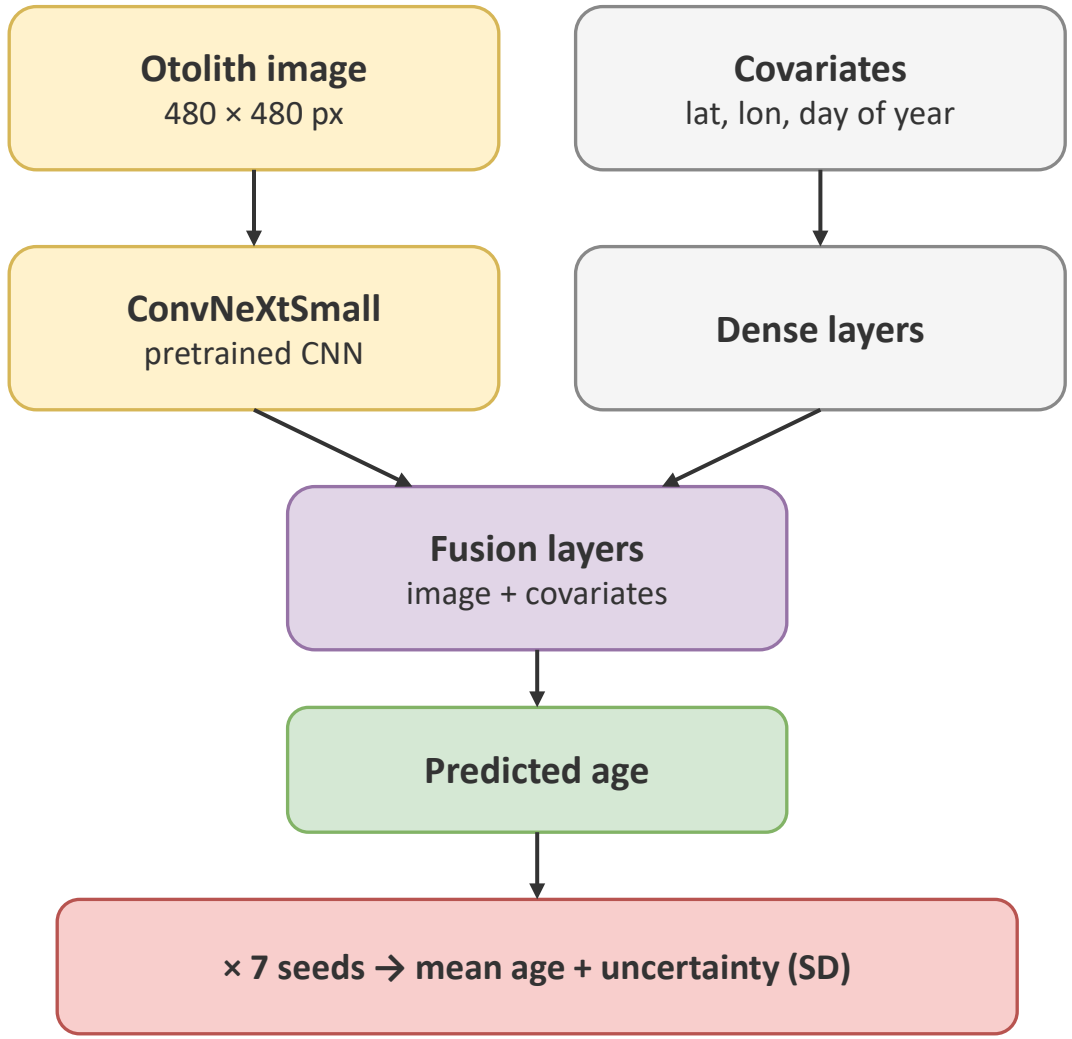
- Pictures are taken with AmScope 8.5MP eyepiece cameras





Modeling approach

- Application of a **convolutional neural network (CNN) model** and **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
- **Transfer learning**: implementation of TensorFlow and Keras libraries, with **ConvNeXtSmall at 480×480 resolution** as preferable configuration
  - earlier testing found alternative architectures underperforming relative to InceptionV3 - with the subsequent expansion of the image database, this conclusion no longer hold
- **Covariate integration**: capture location and day of year enter a parallel branch, fused with image features → modest but consistent gains
- **Deep ensemble** (7 seeds): mean age plus a per-otolith uncertainty (SD across seeds)



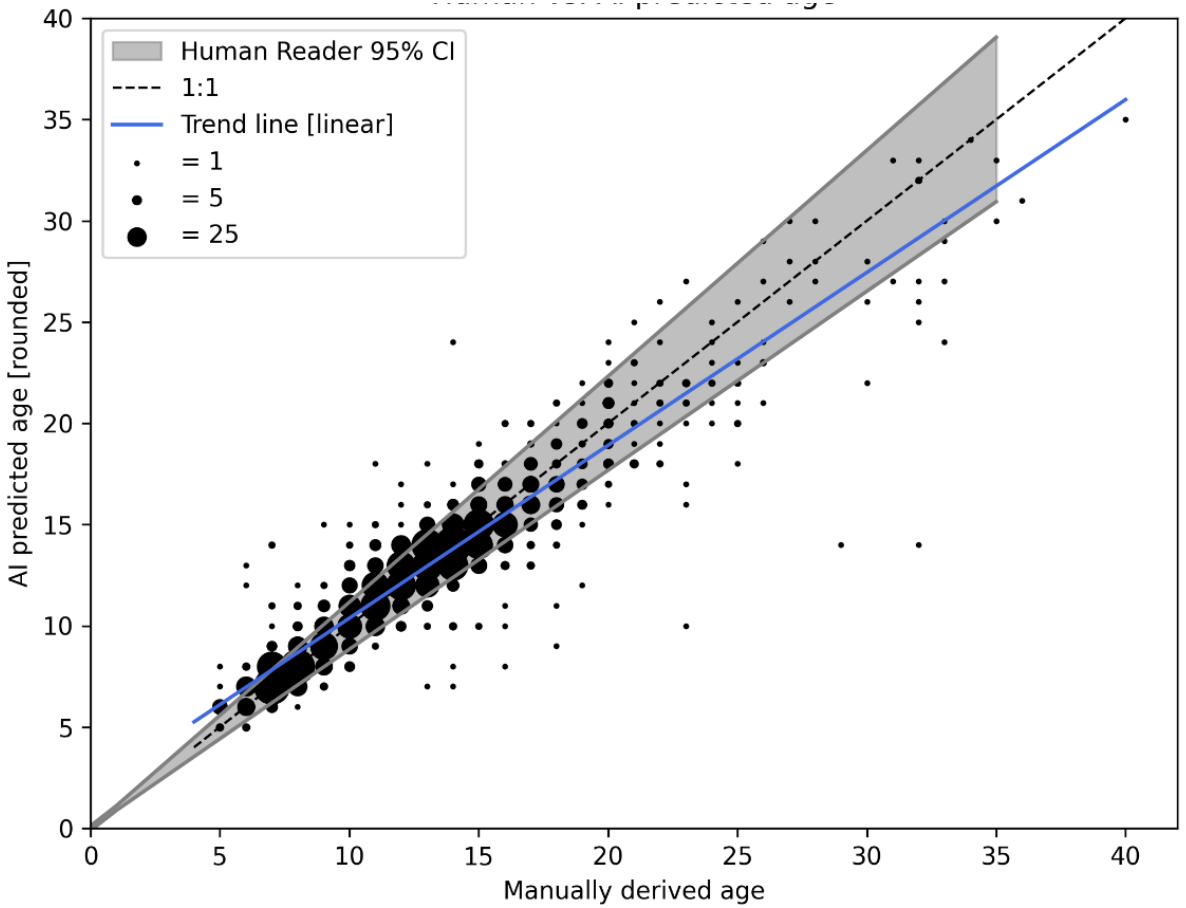
Selected run results

Same images – new architecture

- 11,107 images (7,740 train, 1,367 validation, 2,000 test) – ConvNeXtSmall @ 480 pixels, 7-seed ensemble

The model achieved average RMSE in the test set of **1.62**. On average, the ensemble **correctly predicted age for 36.7% individuals,** plus **43.5% within 1 year**. For reference, human readers agree **60–70%** exactly and **80–90%** within 1 year; **74%** of AI predictions fall inside the readers' 95% CI (grey band)

➢ **Total agreement within 1 year for 80.2% of cases**



Comparison between manually derived age with AI predicted age.

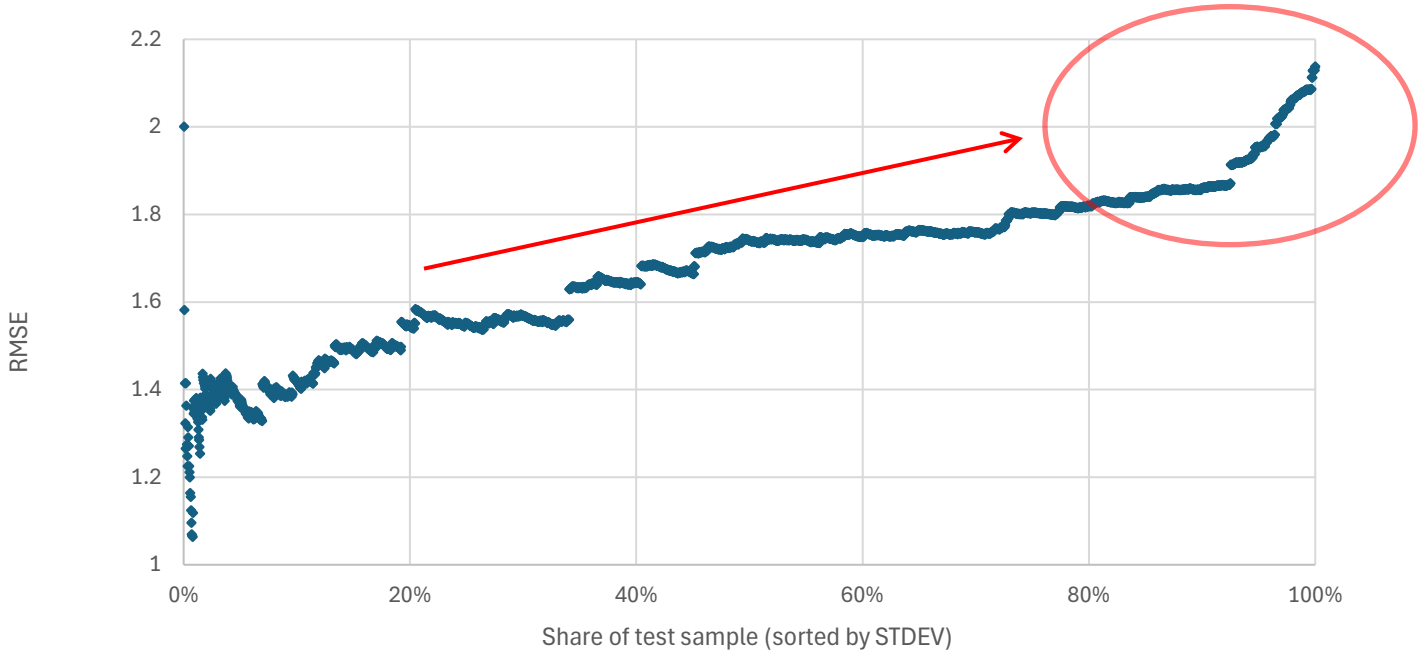
Temporal generalization

- A model trained on 2019 FISS otolith images was evaluated on 2024 data to assess temporal generalization (n = 2,931).
- RMSE increased 1.62 → 2.14 (+32%); predictions within ±1 year 80% → 64%.
- Uncertainty flags hard cases: RMSE 1.82 on the 80% least uncertain otoliths

Fine tuning on 20% of 2024 otoliths (3 seeds), tested on the other 80%:

- Most uncertain 20%: RMSE → 1.75; within ±1 yr → 70%

Uncertainty-guided selection lowers RMSE

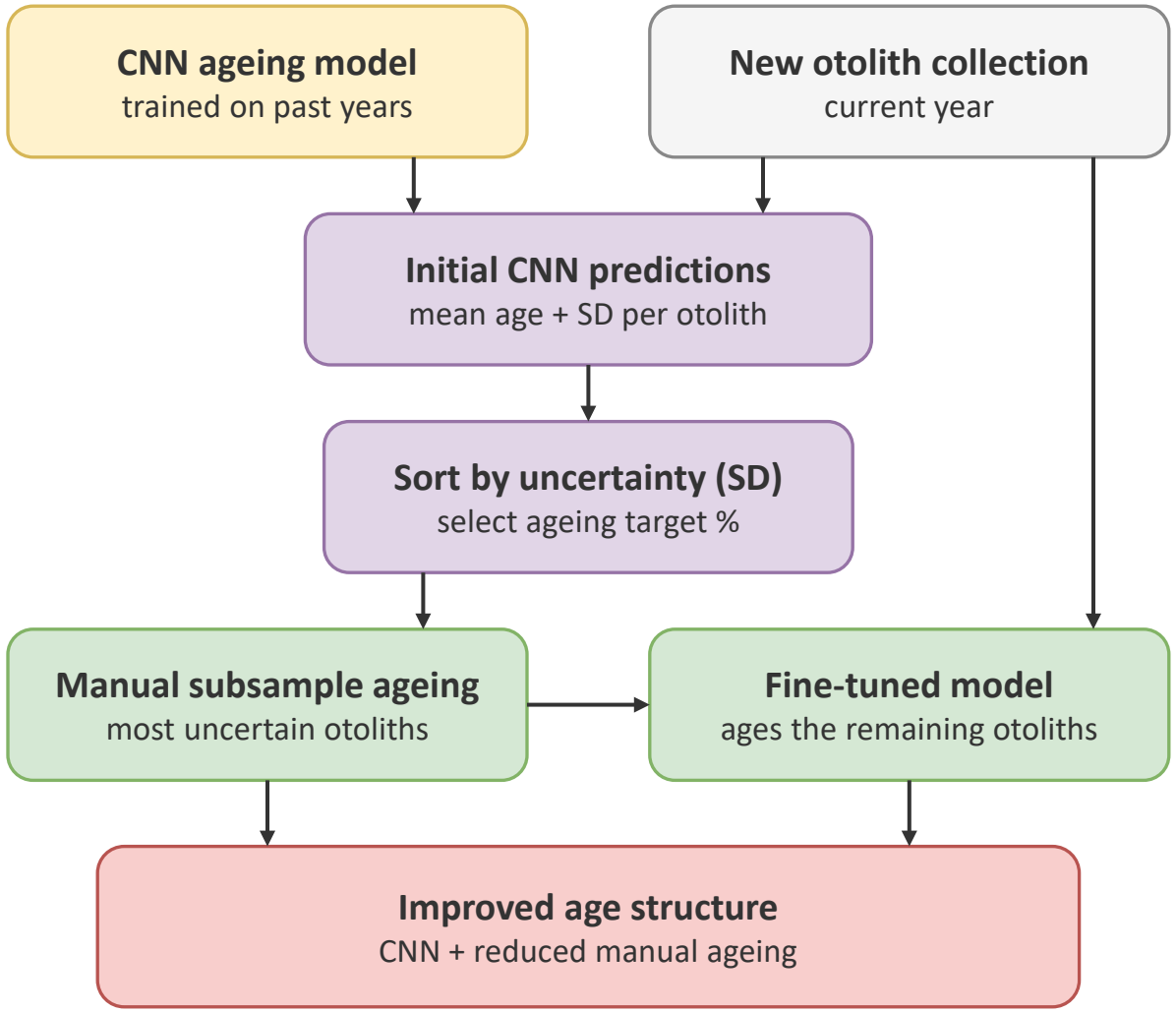


RMSE as a function of cumulative share of the test sample, ordered by prediction uncertainty (standard deviation).

Adopting fine-tuning

Using this in practice:

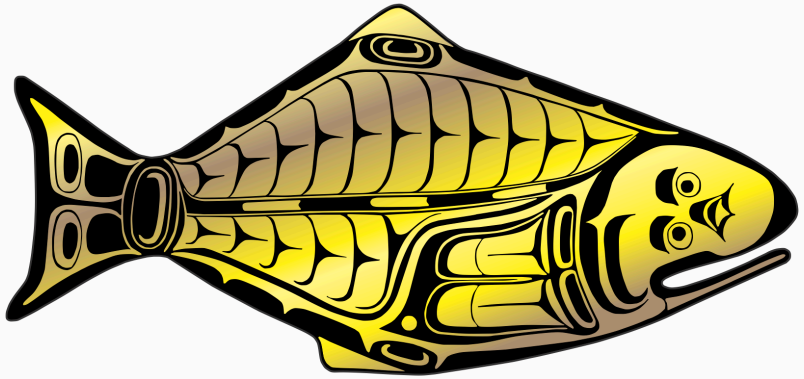
- Most uncertain otoliths go to expert readers
- Model fine-tuned on these new ages
- Training data enriched every year



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
  - 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**.
- **AI is evolving rapidly**, and adapting to new developments may further improve results over time.

INTERNATIONAL PACIFIC



HALIBUT COMMISSION

https://www.iphc.int/