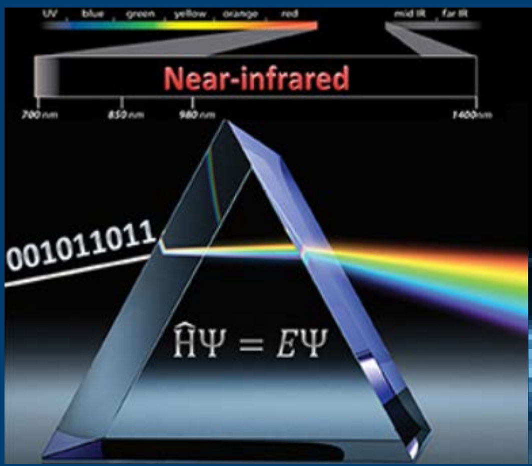
Science, Service, Stewardship




A revolutionary approach for improving age determination efficiency in fish using Fourier transform near infrared spectroscopy (NOAA Fisheries Strategic Initiative)

FT-NIRS 2023 Workshop 3-7 April, Seattle WA

# Envisioning the future of production fish ageing: end-to-end integration of the FT-NIRS age estimation enterprise


*Thomas Helser, Irina Benson, Esther Goldstein, Beth Matta, Brenna Groom, Jon Short*



*Many thanks to AGP staff*


NOAA FISHERIES SERVICE


# Envisioning the future of production fish ageing: end-to-end integration of the FT-NIRS age estimation enterprise

- How does FT-NIRS age prediction work?
- How does FT-NIRS age estimation fit into the TMA process (maintain consistency in both TMA and FT-NIRS)?
- Describe a (***hypothetical***) system of process control and data quality control from end-to-end (reporting examples)
- Emulate the future process: 1) build predictive model on data from past (2014-2018) and predict ages from future collections (2019 and 2021)

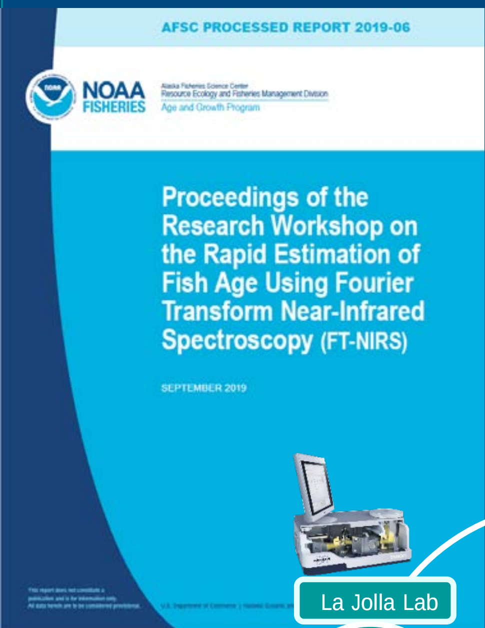
# National Strategic Initiative (R&D) work flow

2020 — 2021 — 2022 — 2023 — 2025 — 2026

**Application Development**

- Instrument optimization (AFSC, Jan-Mar 2019) for otoliths
- FT-NIRS Workshop (April 2019; SI planning over 5 years)
- Otolith spectra acquisition (3 species per region x 5-year time depth
- Predictive model development (calibration/validation)

Discovery switchback


AFSC PROCESSED REPORT 2019-06

NOAA FISHERIES

Alaska Fisheries Science Center
Resource Ecology and Fisheries Management Division
Age and Growth Program

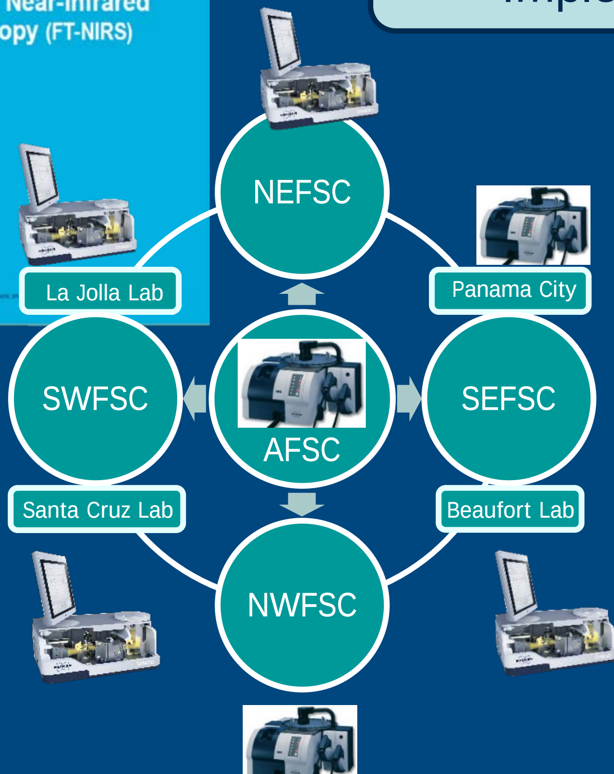
Proceedings of the Research Workshop on the Rapid Estimation of Fish Age Using Fourier Transform Near-Infrared Spectroscopy (FT-NIRS)

SEPTEMBER 2019

**Application Implementation**

- Process control, quality control, fault detection
- Standards, best practices (simulation)
- AI/Deep machine learning
- Build scientific basis of tech. (publish)

Discovery switchback



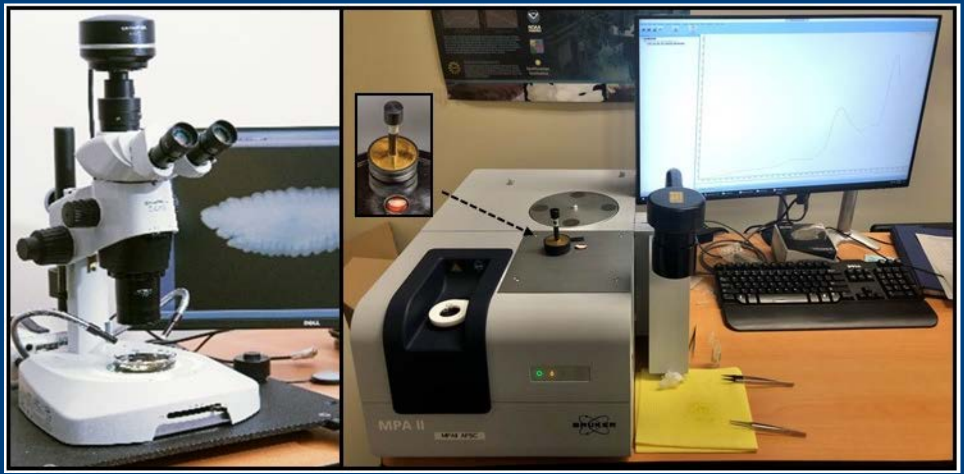

**Stock Assessment Integration**

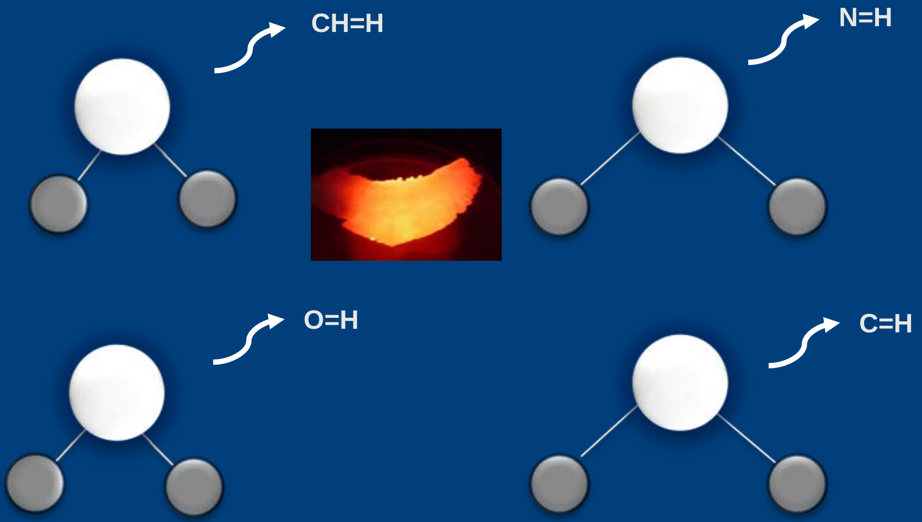
- Evaluation of assessment model outputs to FT-NIRS data
- Provide FT-NIRS precision & reliability metrics

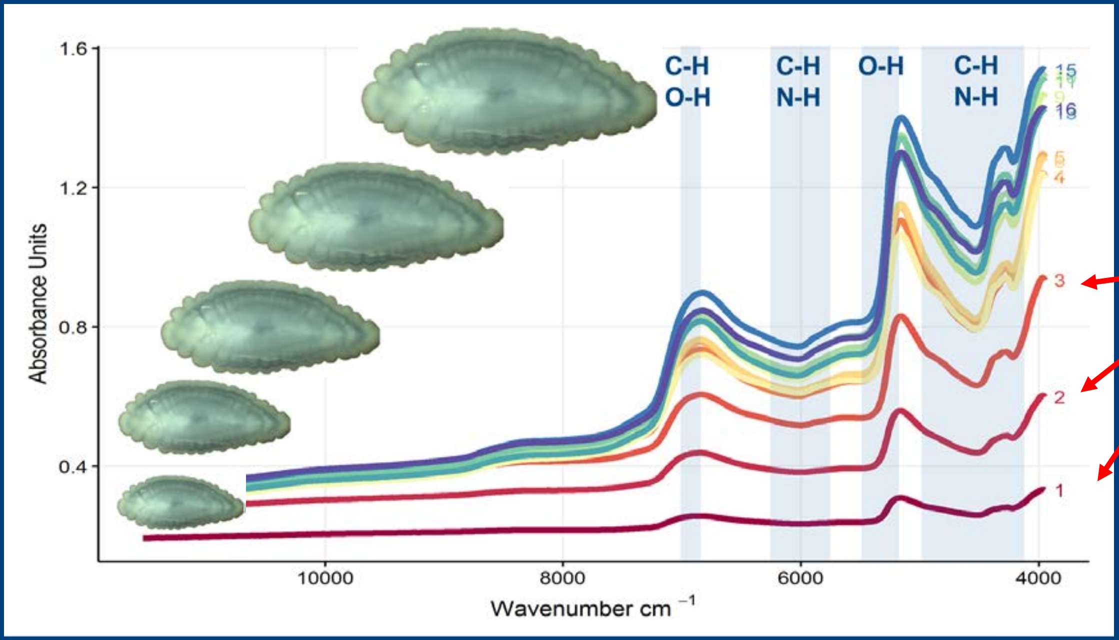
**Deployment of technology**

- Integrating technology into current production setting (species-specific)

# NIR Spectroscopy: measurement of intra-molecular vibrations





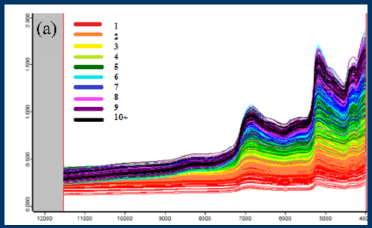

Different ages have different absorbance profiles

# How is fish age predicted from otolith spectra?

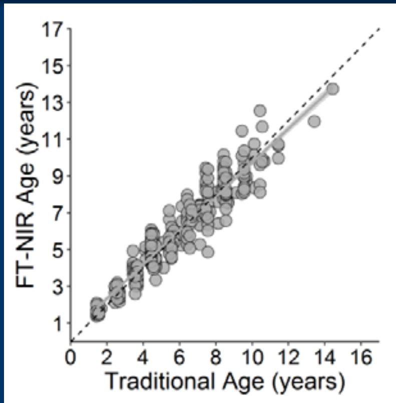

Samples

Measured by FT-NIR

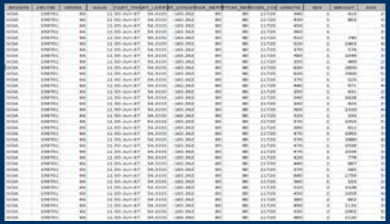
Spectral Data




Deep machine learning




Analyzed by Primary Reference Method

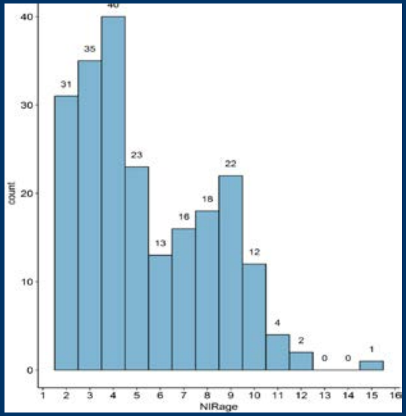


Reference Values (ages)

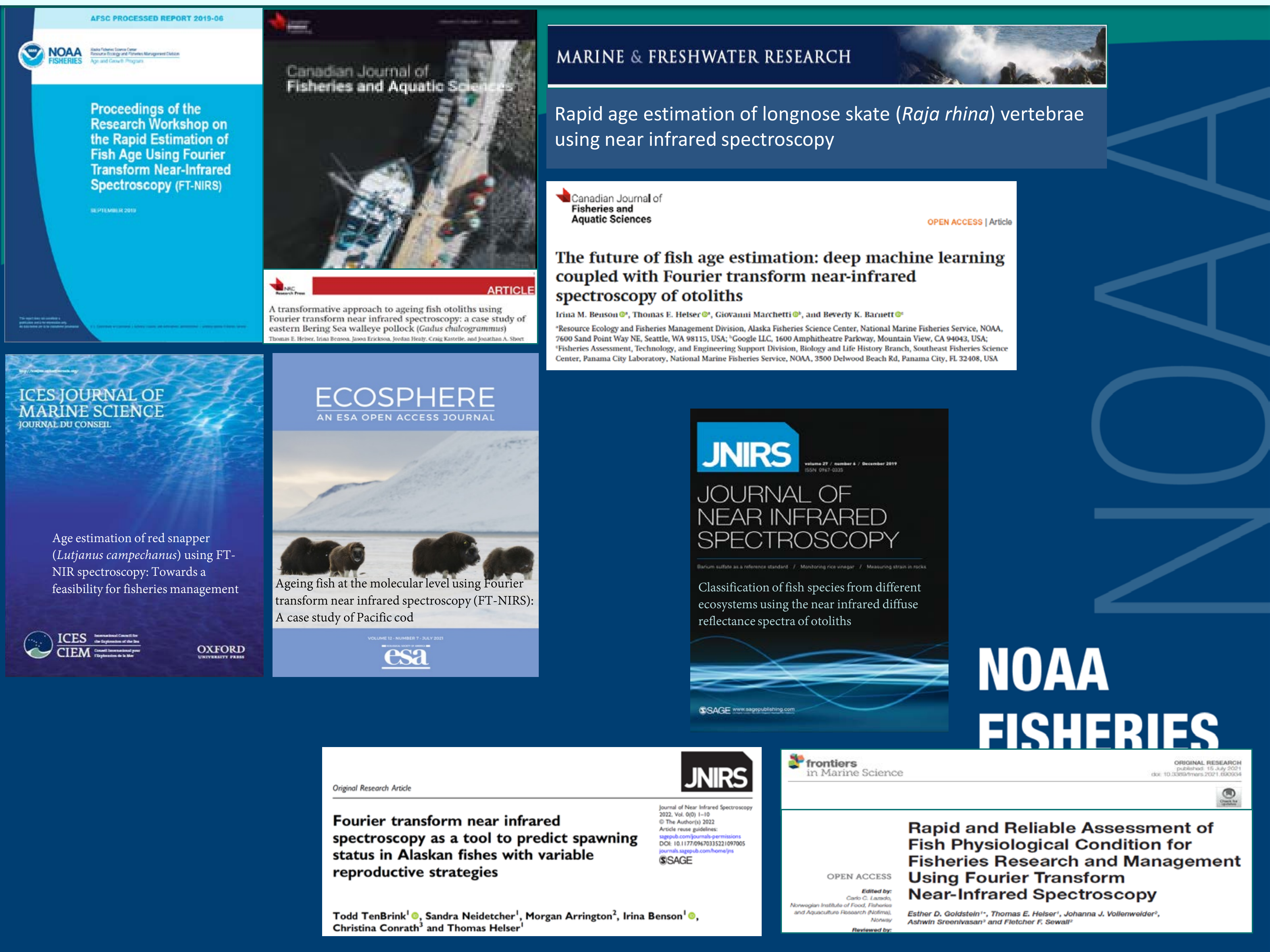
Model uploaded to instrument

Unknown Samples

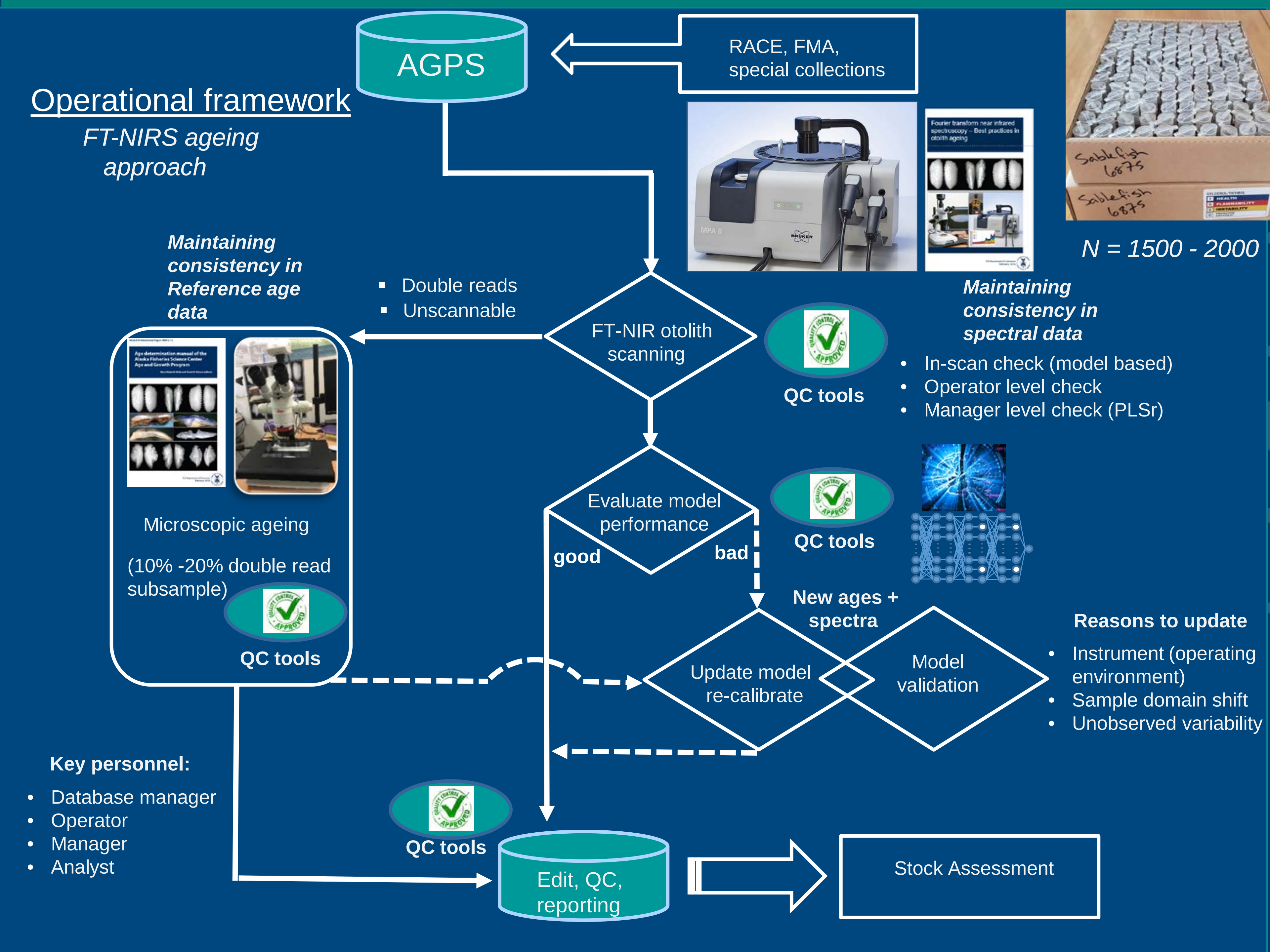
Measured by FT-NIR




Age compositions for assessments



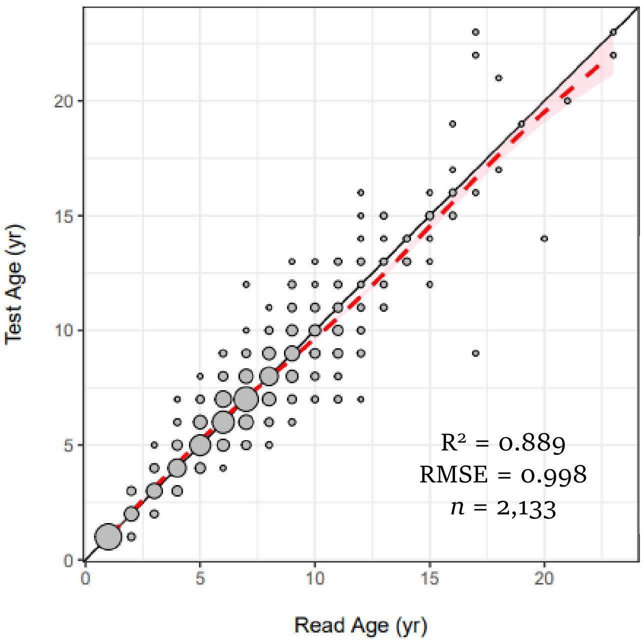
AFSC PROCESSED REPORT 2019-06
NOAA FISHERIES
Proceedings of the Research Workshop on the Rapid Estimation of Fish Age Using Fourier Transform Near-Infrared Spectroscopy (FT-NIRS)
SEPTEMBER 2019
Canadian Journal of Fisheries and Aquatic Sciences
ARTICLE
A transformative approach to ageing fish otoliths using Fourier transform near infrared spectroscopy: a case study of eastern Bering Sea walleye pollock (Gadus chalcogrammus)
Thomas E. Helser, Irina Benson, Jason Erickson, Jordan Healy, Craig Kastelle, and Jonathan A. Short
MARINE & FRESHWATER RESEARCH
Rapid age estimation of longnose skate (Raja rhina) vertebrae using near infrared spectroscopy
Canadian Journal of Fisheries and Aquatic Sciences
OPEN ACCESS | Article
The future of fish age estimation: deep machine learning coupled with Fourier transform near-infrared spectroscopy of otoliths
Irina M. Benson, Thomas E. Helser, Giovanni Marchetti, and Beverly K. Barnett
Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, WA 98115, USA; Google LLC, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA; Fisheries Assessment, Technology, and Engineering Support Division, Biology and Life History Branch, Southeast Fisheries Science Center, Panama City Laboratory, National Marine Fisheries Service, NOAA, 3500 Delwood Beach Rd, Panama City, FL 32408, USA
ICES JOURNAL OF MARINE SCIENCE
JOURNAL DU CONSEIL
Age estimation of red snapper (Lutjanus campechanus) using FT-NIR spectroscopy: Towards a feasibility for fisheries management
ICES CIEM
OXFORD UNIVERSITY PRESS
ECOSPHERE
AN ESA OPEN ACCESS JOURNAL
Ageing fish at the molecular level using Fourier transform near infrared spectroscopy (FT-NIRS): A case study of Pacific cod
VOLUME 12 • NUMBER 7 • JULY 2021
esa
JNIRS
volume 27 / number 6 / December 2019
JOURNAL OF NEAR INFRARED SPECTROSCOPY
Classification of fish species from different ecosystems using the near infrared diffuse reflectance spectra of otoliths
SAGE www.sagepublishing.com
NOAA FISHERIES
JNIRS
Original Research Article
Fourier transform near infrared spectroscopy as a tool to predict spawning status in Alaskan fishes with variable reproductive strategies
Journal of Near Infrared Spectroscopy 2022, Vol. 0(0) 1–10
© The Author(s) 2022
SAGE
Todd TenBrink, Sandra Neidetcher, Morgan Arrington, Irina Benson, Christina Conrath and Thomas Helser
frontiers in Marine Science
ORIGINAL RESEARCH
OPEN ACCESS
Rapid and Reliable Assessment of Fish Physiological Condition for Fisheries Research and Management Using Fourier Transform Near-Infrared Spectroscopy
Esther D. Goldstein, Thomas E. Helser, Johanna J. Vollenweider, Ashwin Sreenivasan and Fletcher F. Sewall
NOAA



Operational framework
FT-NIRS ageing approach
AGPS
RACE, FMA, special collections
N = 1500 - 2000
Maintaining consistency in Reference age data
Double reads
Unscannable
FT-NIR otolith scanning
QC tools
Maintaining consistency in spectral data
In-scan check (model based)
Operator level check
Manager level check (PLSr)
Microscopic ageing
(10% -20% double read subsample)
QC tools
Evaluate model performance
good
bad
QC tools
New ages + spectra
Update model re-calibrate
Model validation
Reasons to update
Instrument (operating environment)
Sample domain shift
Unobserved variability
Key personnel:
Database manager
Operator
Manager
Analyst
QC tools
Edit, QC, reporting
Stock Assessment

- **Using otoliths from past EBS 2014-2018 surveys build a predictive CNN model.**
- **Predict ages for 2019 & 2021 using FT-NIRS (2014 – 2018) model**

# EBS walleye pollock (2014-2018) Age data



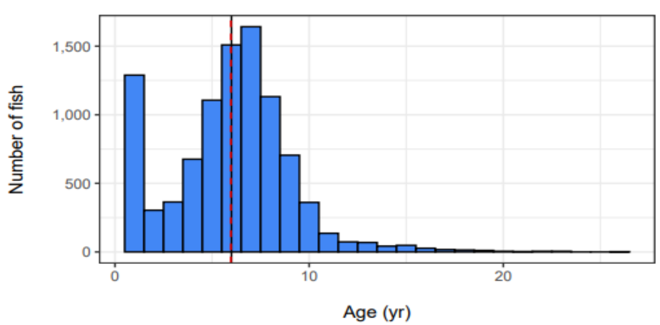

| Precision statistic | Value |
|---|---|
| Percent agreement (PA) | 66.9% |
| Average percent error (APE) | 3.05% |
| Coefficient of variation (CV) | 4.32% |
| Total number of fish in ageing collection | 9603 |
| Number of fish unaged | 58 |
| Number of fish in precision-testing sample | 2133 |
| Number of fish aged by two readers | 2133 |
| Percentage of fish with paired age readings | 22 |
| | 5.83 |

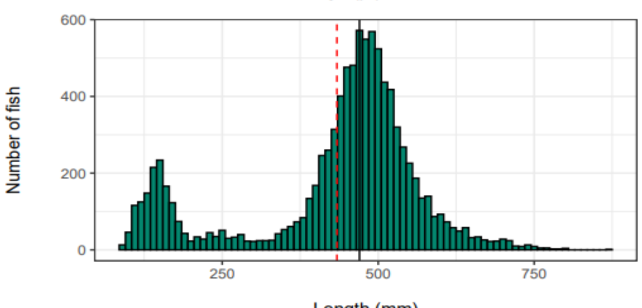
(a) Bias direction

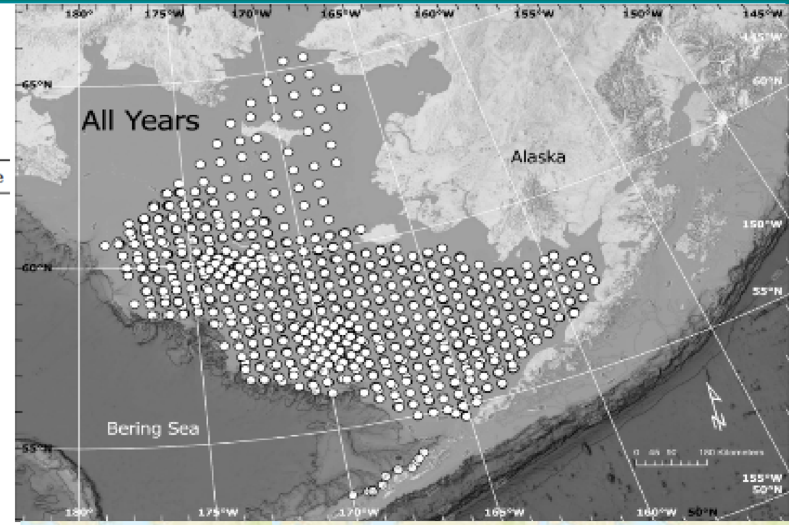
| | | |
|---|---|---|
| minus bias | 371 otoliths | 17.4% |
| plus bias | 335 otoliths | 15.7% |
| complete agreement | 1427 otoliths | 66.9% |

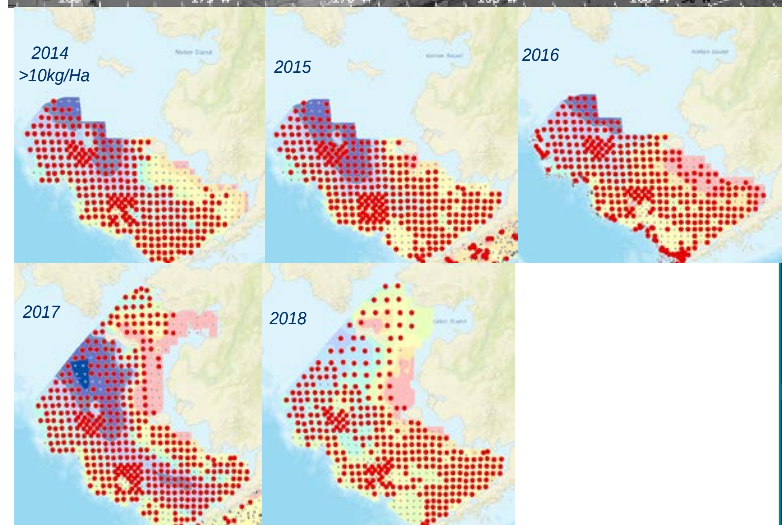
(b) Tests of symmetry

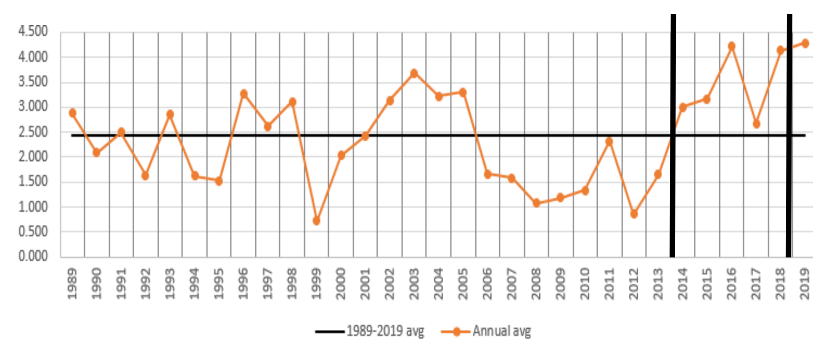
| Test name | df | Test statistic | *p* |
|---|---|---|---|
| Bowker's | 48 | 55.72 | 0.21 |
| Evans-Hoenig | 7 | 4.79 | 0.69 |



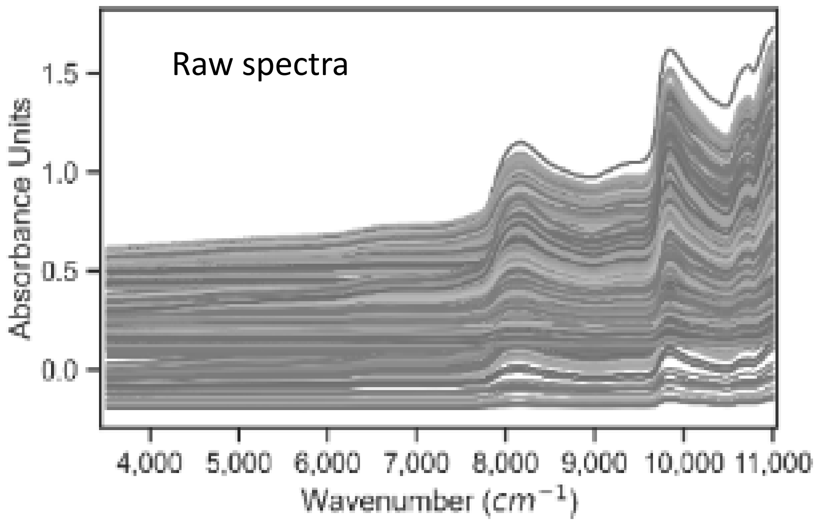



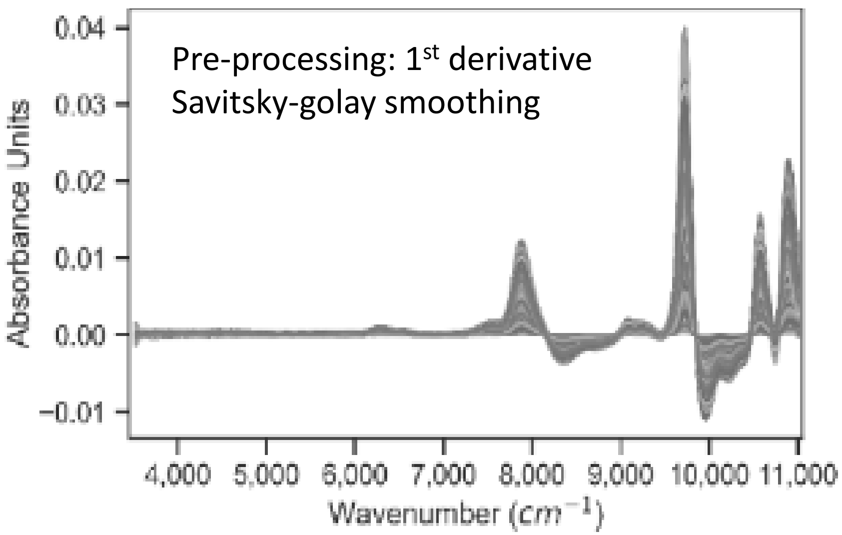



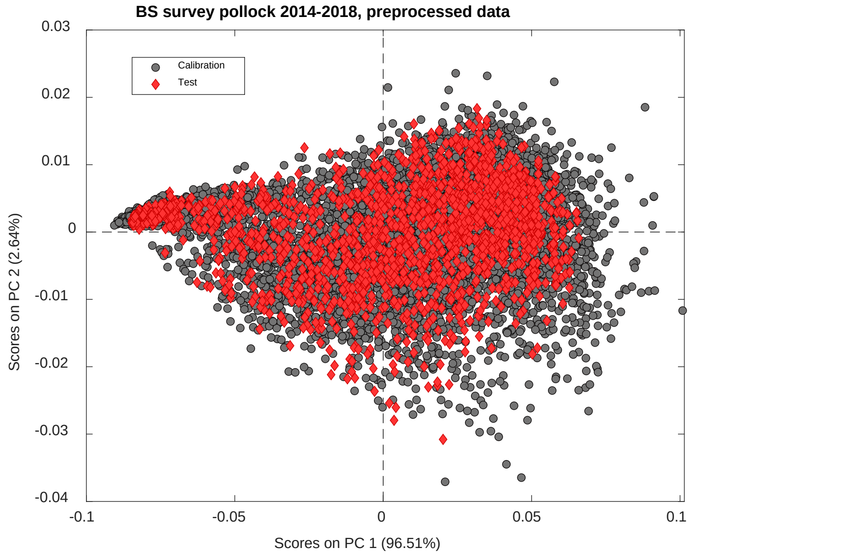



# EBS walleye pollock base model (2014-2018) Otolith spectra

- Onion algorithm to split train & test (applied to each year)



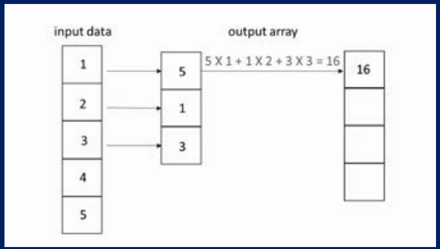




| Features | Subset | Values |
|---|---|---|
| **Mean fish age* (standard deviation)** | train | 5.72 (3.3) |
| | test | 5.82 (3.14) |
| **Fish age range*** | train | 1 – 23 |
| | test | 1 – 23 |
| **Fish length range** | train | 90 – 870 |
| | test | 100 – 800 |
| **Latitude range** | train | 53.13822 – 65.25117 |
| | test | 53.15808 – 65.23351 |
| **Gear depth range** | train | 18 – 695.79 |
| | test | 18 – 640 |
| **Gear temperature range** | train | -1.6 – 9.9 |
| | test | -1.6 – 9.6 |

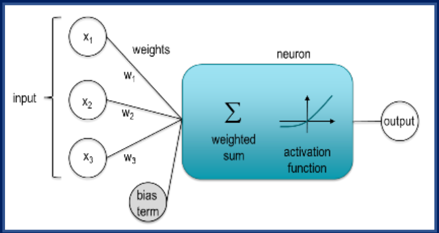
# Convolutional Neural Network (Multi Mode)

Convolutional layer consists of a kernel that slides along our data and applies its weights to the data values.

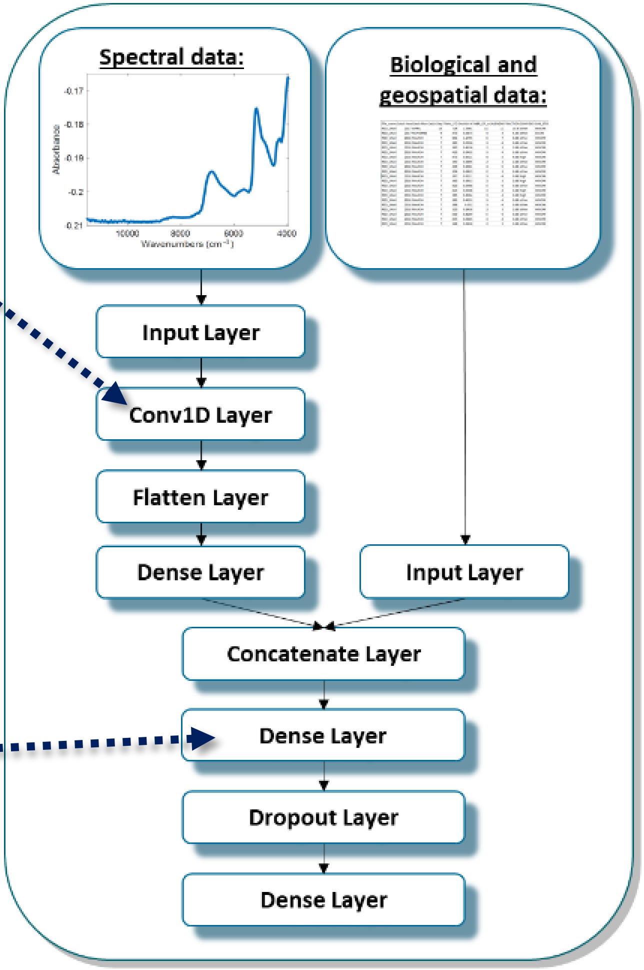



Deep learning networks will have multiple kernels and will produce multiple output arrays.

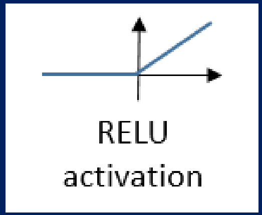
Neurons are core processing units of the network.




Dense layers of neural network is made up of layers of neurons.

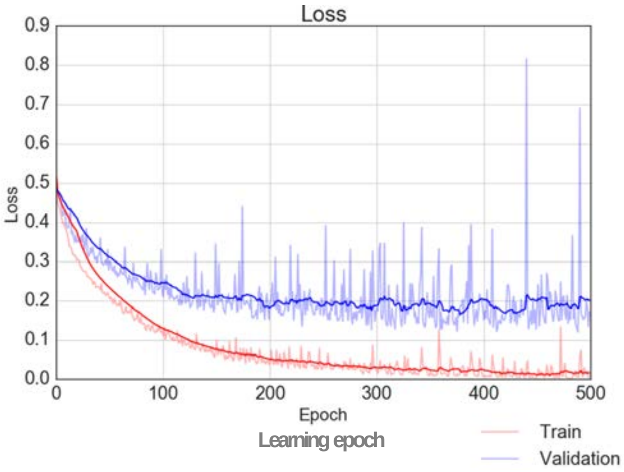



Non-linear activation functions introduce non-linear properties into network.




We used Rectified linear unit (RELU) functions which outputs the input directly if it is positive, for negative input outputs zero.

To implement our models we employed Python using TensorFlow with Keras API and hyperband optimization (HB) for hyperparameter tuning.



# Predictive results from CNN

Walleye pollock
(*Gadus chalcogrammus*)

Canadian Journal of Fisheries and Aquatic Sciences

OPEN ACCESS | Article

**The future of fish age estimation: deep machine learning coupled with Fourier transform near-infrared spectroscopy of otoliths**

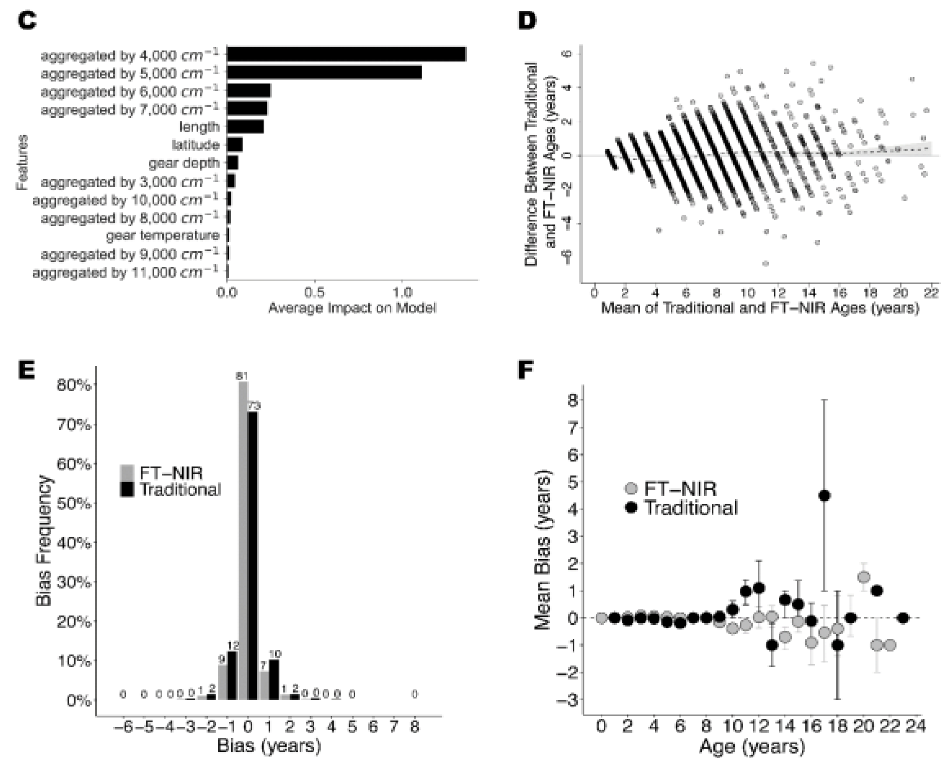
Irina M. Benson[a], Thomas E. Helser[a], Giovanni Marchetti[b], and Beverly K. Barnett[c]

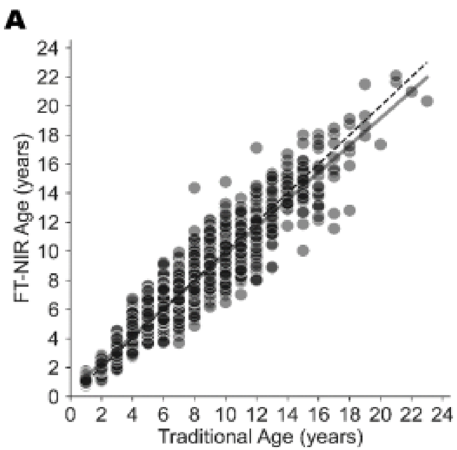
[a]Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, WA 98115, USA; [b]Google LLC, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA; [c]Fisheries Assessment, Technology, and Engineering Support Division, Biology and Life History Branch, Southeast Fisheries Science Center, Panama City Laboratory, National Marine Fisheries Service, NOAA, 3500 Delwood Beach Rd, Panama City, FL 32408, USA

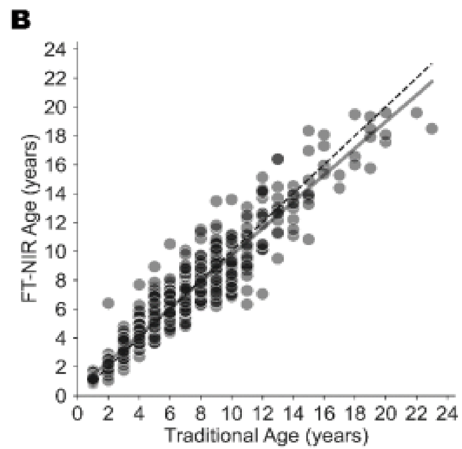
Table 3. Prediction results of multi-modal CNN and PLS models.

| Model | Number of otoliths | | | $R^2$ | | RMSE | |
|---|---|---|---|---|---|---|---|
| | Train | Test | Outliers* | Train | Test | Train | Test |
| CNN | | | | **0.93** | **0.92** | **0.83** | **0.91** |
| PLS with all spectral wavenumbers | 6866 | 1751 | 12 | 0.89 | 0.87 | 0.99 | 1.14 |
| PLS with selected spectral wavenumbers | | | | 0.90 | 0.87 | 0.97 | 1.12 |

*Spectral outliers removed for each collection year:
2014 (n=3); 2015 (n=1); 2016 (n=3); 2017 (n=1); 2018 (n=4)



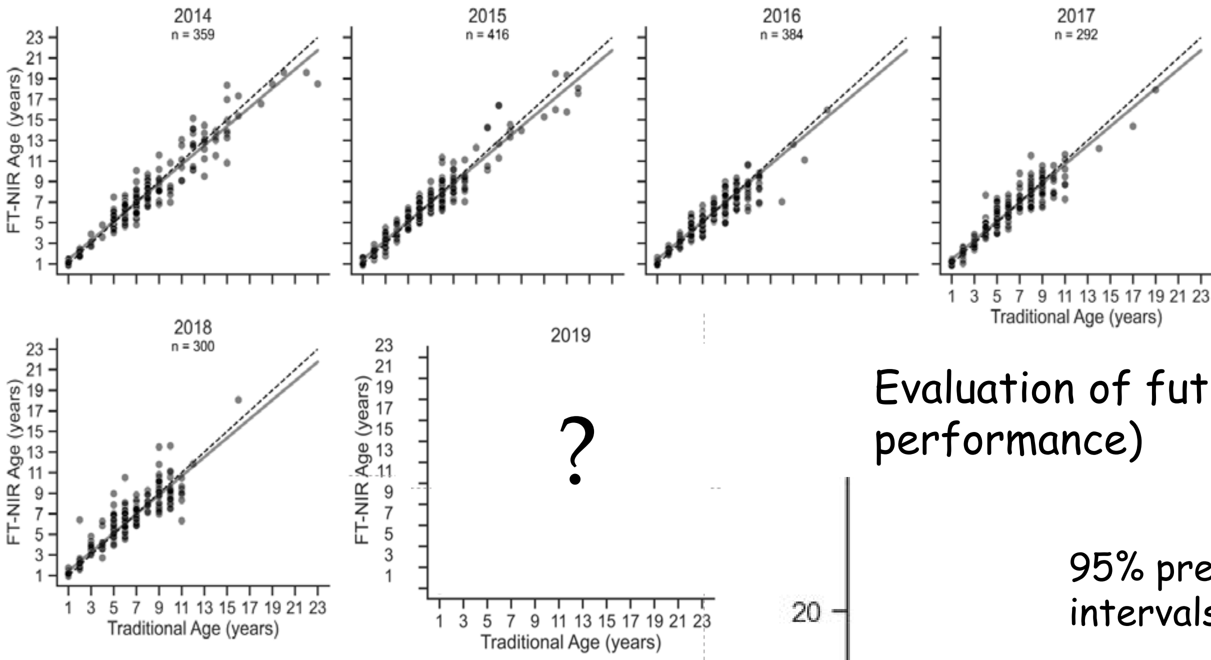




$$B_R TRAD = (Age^{Reader1} - Age^{Reader2})$$

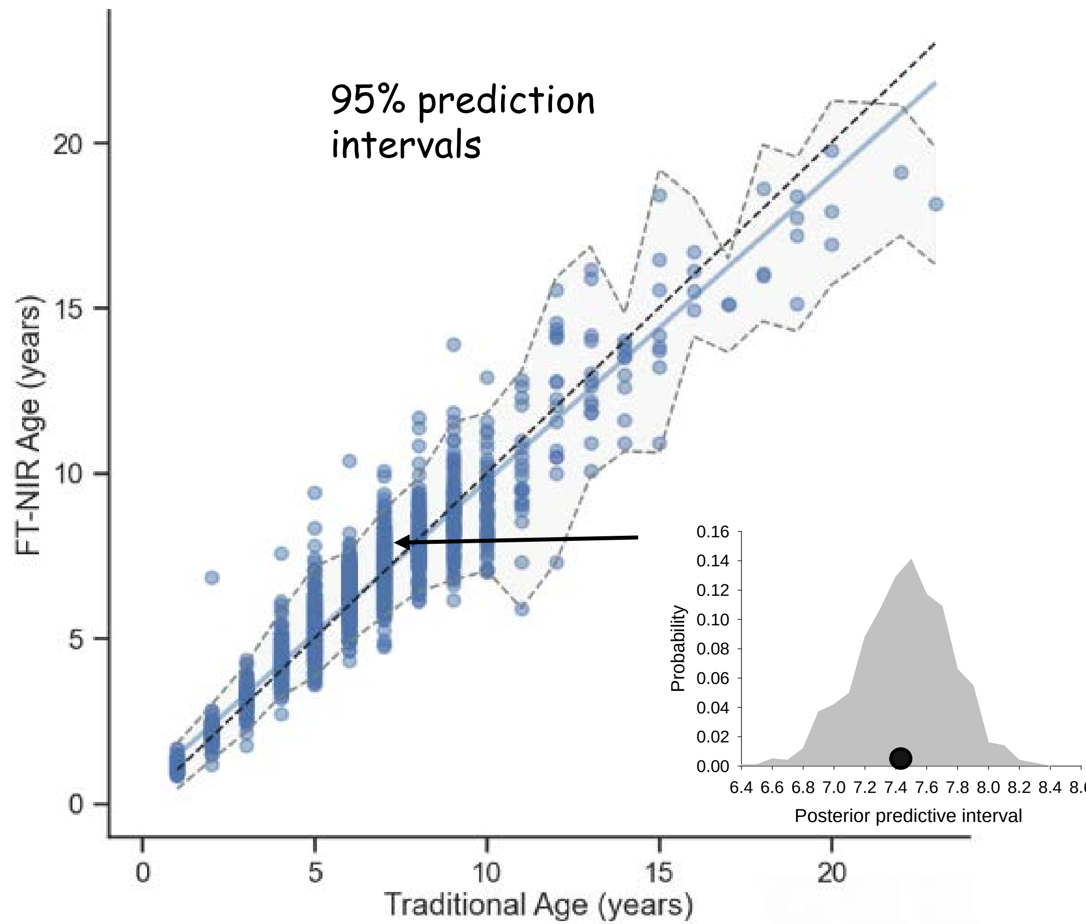
$$B_R FT - NIRS = (Age^{FT-NIRS} - Age^{Final\ age})$$

FT-NIRS
$R^2 = 0.92$
$RMSE_p = 0.912$

TMA
$R^2 = 0.889$
RMSE = 0.998




## Evaluation of future predictions (model performance)




**Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning**

*Proceedings of the 33rd International Conference on Machine Learning*, New York, NY, USA, 2016. JMLR: W&CP volume 48. Copyright 2016 by the author(s).

nature > nature reviews physics > viewpoint > article

Viewpoint | Published: 22 August 2022

**Bayesian uncertainty quantification for machine-learned models in physics**

Yarin Gal, Petros Koumoutsakos, Francois Lanusse, Gilles Louppe & Costas Papadimitriou

*Nature Reviews Physics* 4, 573–577 (2022) | Cite this article

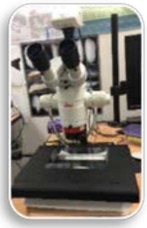
4138 Accesses | 23 Altmetric | Metrics

## Collection of pollock otolith NIR spectra
## (*scanning 2019 & 2021 EBS survey collections*)

## Processing 2019 & 2021 age data (only ageing 20% of entire collection)

*2019*

***Consistency in Reference age data (assumes 20% double reads)***



**QC tools**

Table 1: Precision statistics for the Walleye Pollock 2019 Bering Sea Survey collection.

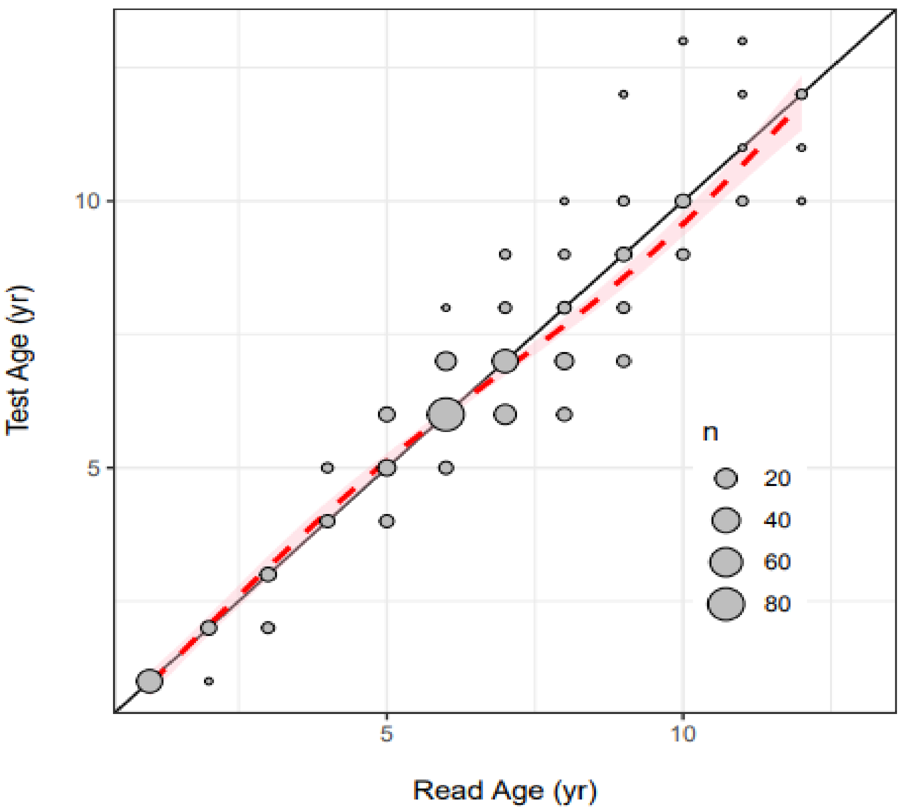
| Precision statistic | Value |
|---|---|
| Percent agreement (PA) | 64.61% |
| Average percent error (APE) | 3.2% |
| Coefficient of variation (CV) | 4.53% |
| Total number of fish in ageing collection | 1552 |
| Number of fish unaged | 13 |
| Number of fish in precision-testing sample | 308 |
| Number of fish aged by two readers | 308 |
| Percentage of fish with paired age readings | 20 |
| Average read age (paired reads only) | 5.95 |
| Average test age (paired reads only) | 5.86 |

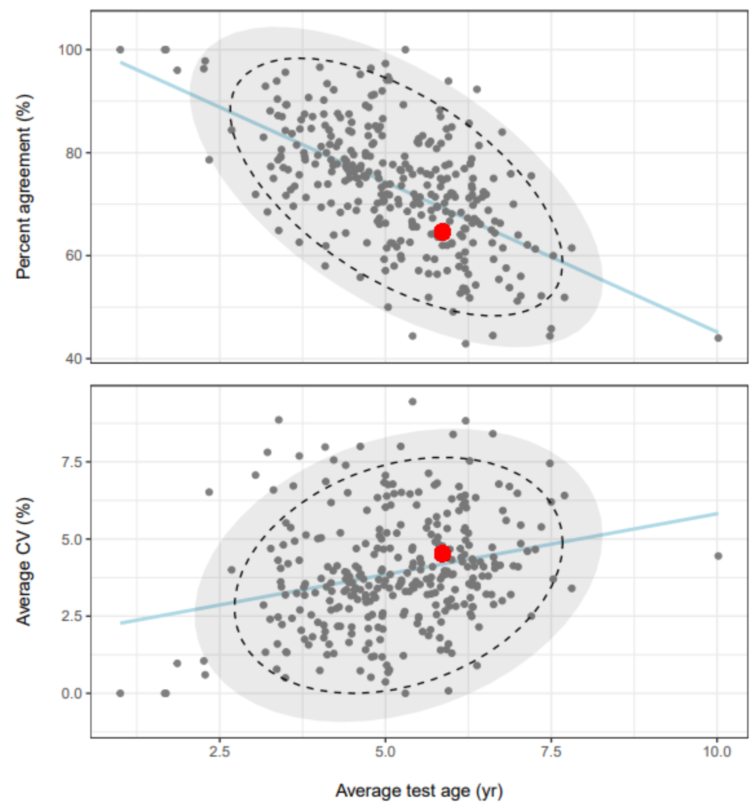
(a) Bias direction

| | | |
|---|---|---|
| minus bias | 42 otoliths | 13.6% |
| plus bias | 67 otoliths | 21.8% |

(b) Tests of symmetry

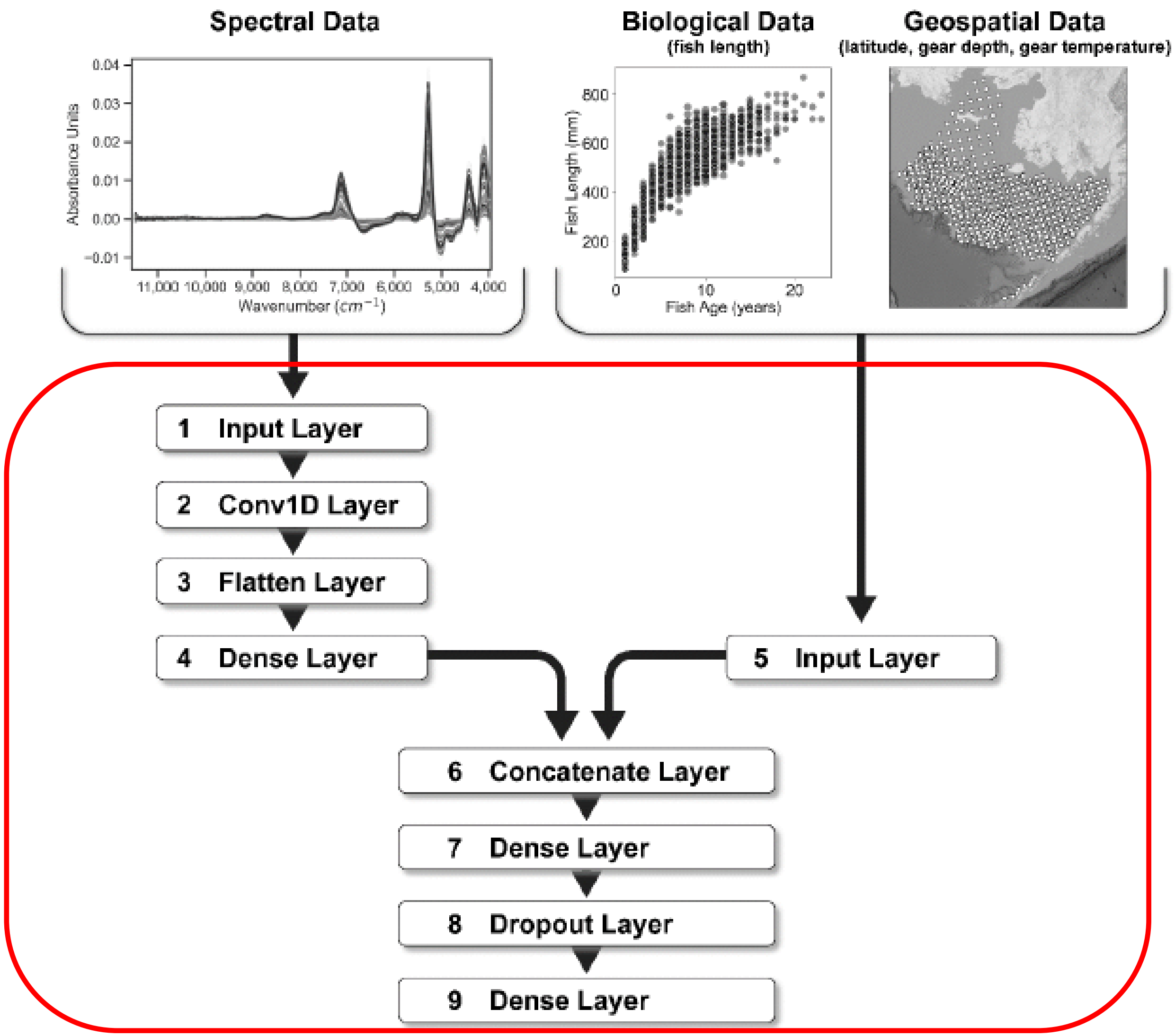
| Test name | df | Test statistic | *p* |
|---|---|---|---|
| Bowker's | 17 | 23.44 | 0.14 |
| Evans-Hoenig | 3 | 9.1 | 0.03 |

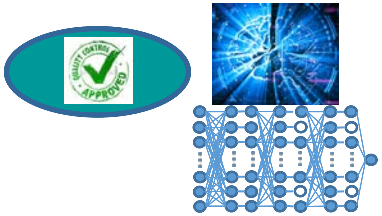


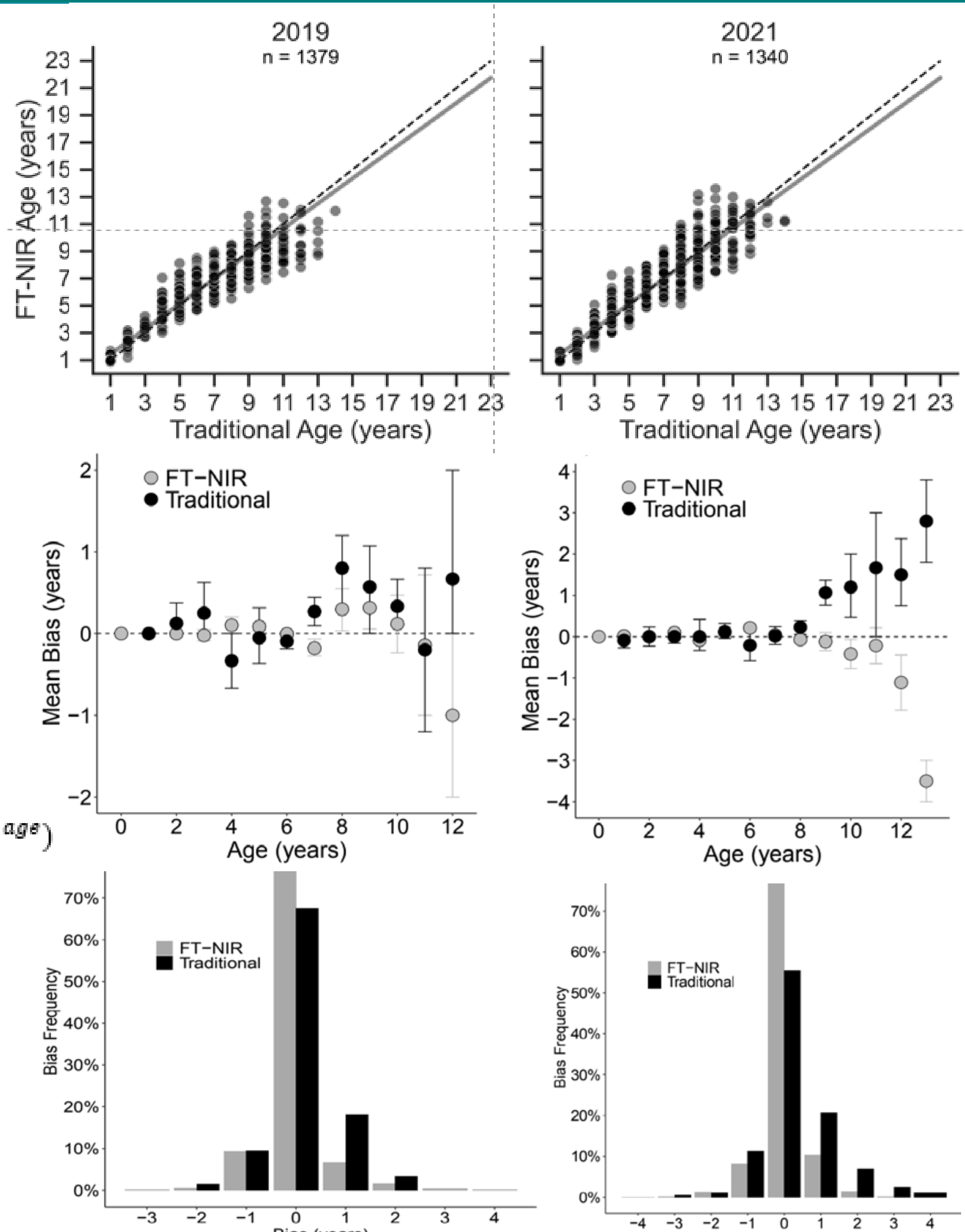



# Predicting 2019 & 2021 (*ages, spectra & metadata*)

**2014-2018 CNN base model**



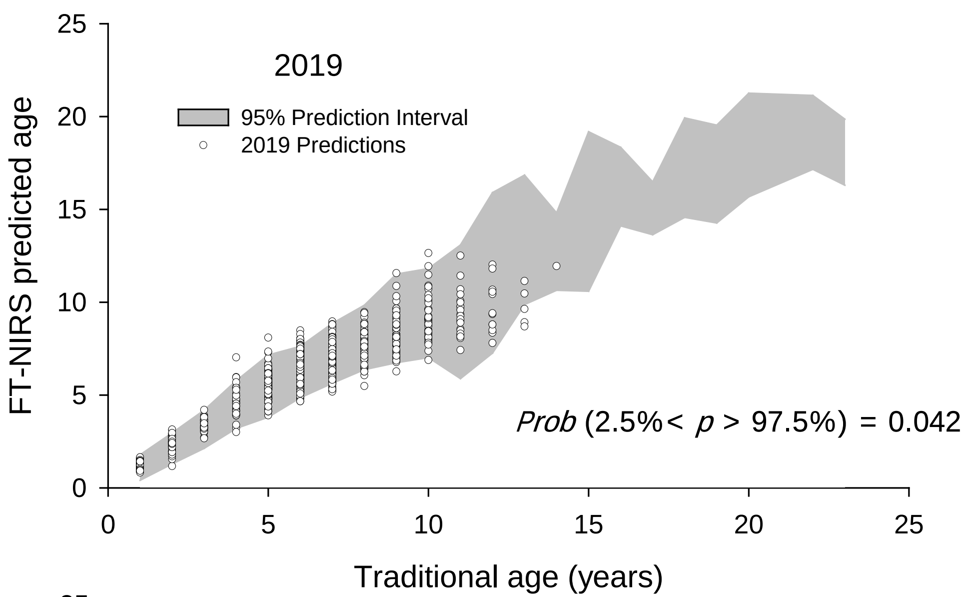


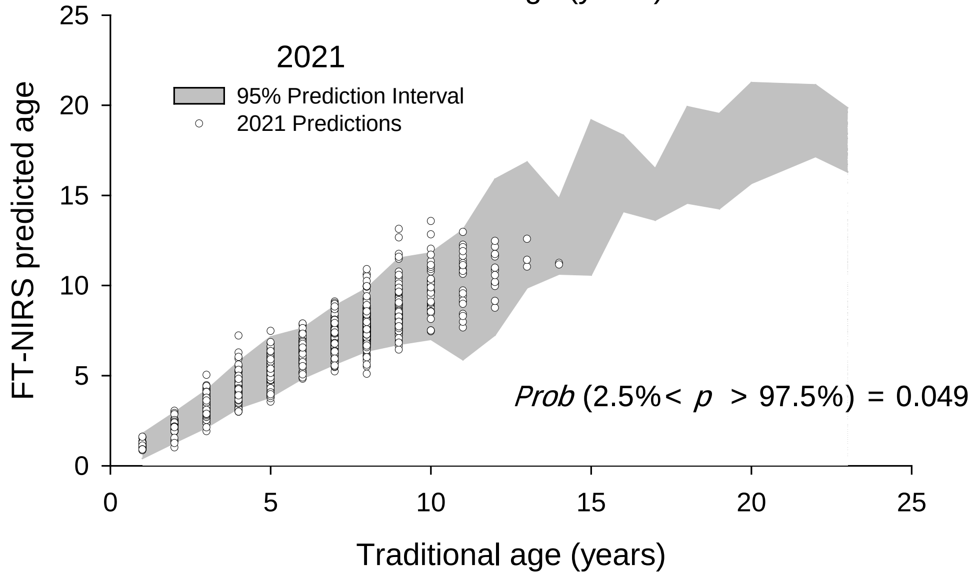



$$B_R TRAD = (Age^{Reader1} - Age^{Reader2})$$

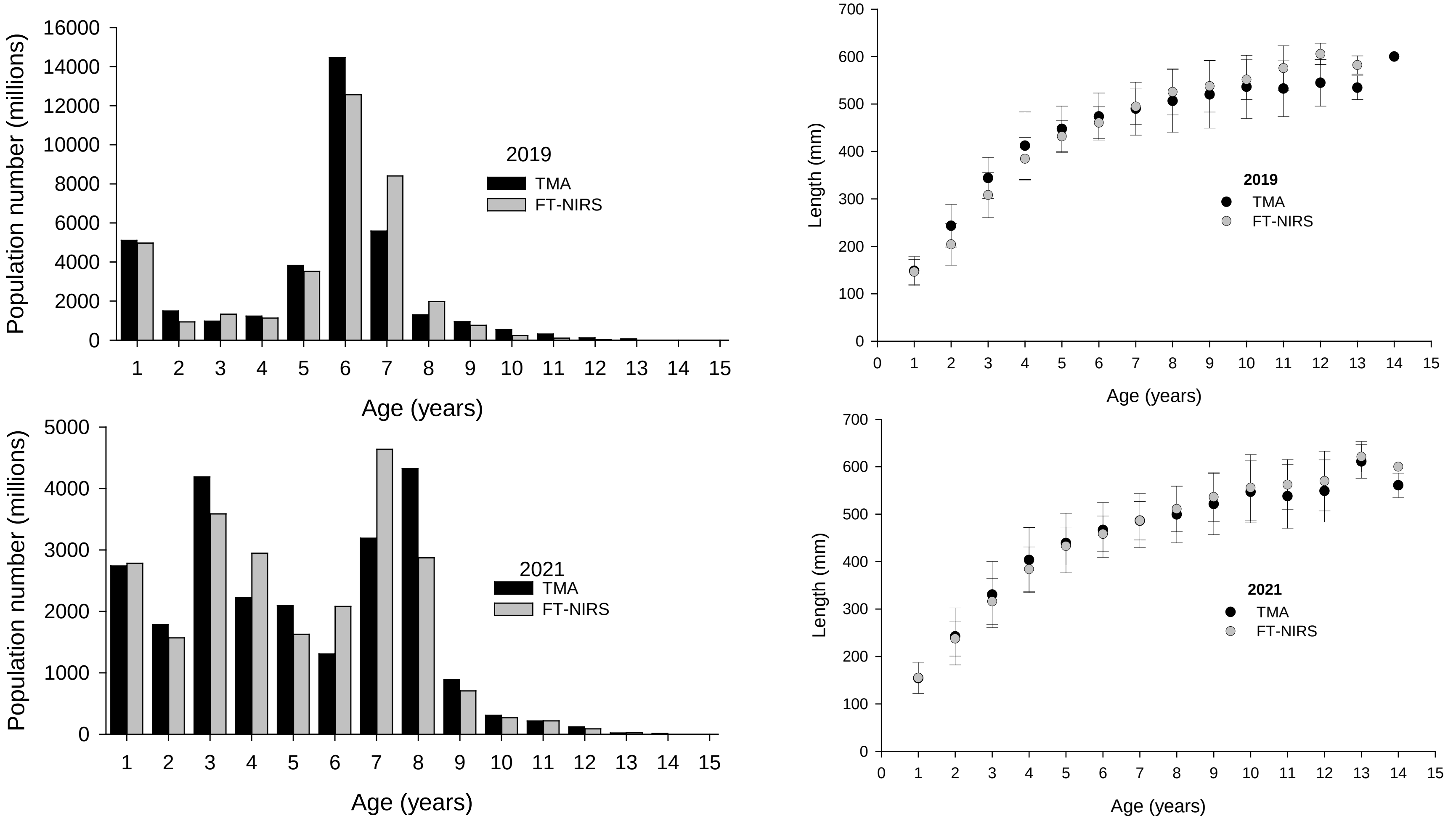
$$B_R FT-NIRS = (Age^{FT-NIRS} - Age^{Final\ age})$$

# Assessing CNN model performance of future predictions



# Age data products for stock assessments




Envisioning the future of production fish ageing: end-to-end integration of the FT-NIRS age estimation enterprise

## *Talking points?*

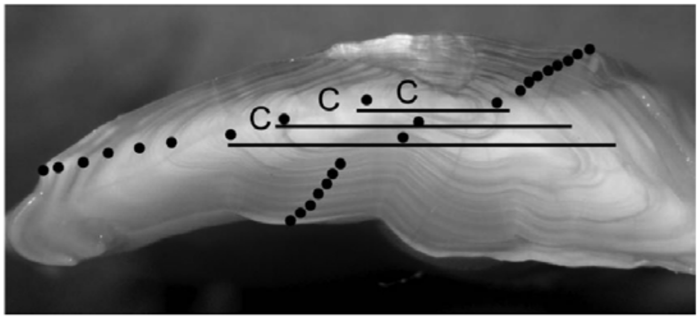
- Define the “bar” for successful application
- Ageing imprecision for use in stock assessments (TMA + NIR)
- FT-NIRS age data performance in stock assessments – model updating to accommodate unseen variability
- FT-NIRS operational transition and technological deployment
- Utility function – evaluate trade offs between TMA samples (double reads + outliers + issue otoliths) and FT-NIRS efficiency gains


The Future:

- Continue to investigate ageing error and bias effects on models (enhance predictive models to accommodate “known” age fish)
- Improve database interface and architecture (employ time-flow statistics for cost-benefit analysis)
- Broaden simulation framework to accommodate larger range of species & life histories
- Leverage cloud computing and machine learning (take advantage of other data types)
- Develop predictive model tool box (R, ADMB, OPUS, Python) - standardization
- Better define staff needs and skill sets required for future operationalization
- Think more about operational transition and technological deployment (where possible)
- Communicate to stake holders

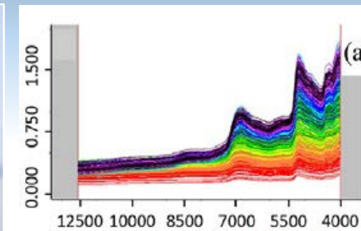
Questions?

# Pacific cod otolith (transverse section)

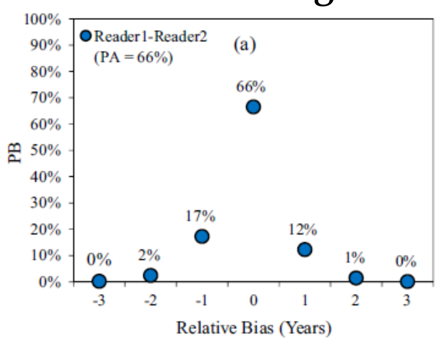


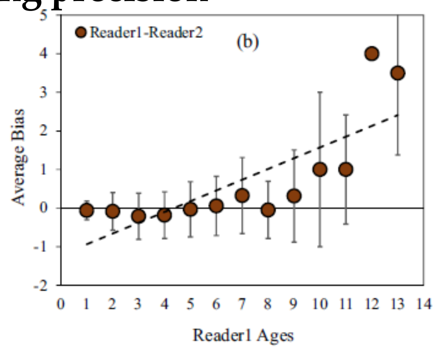

COASTAL AND MARINE ECOLOGY

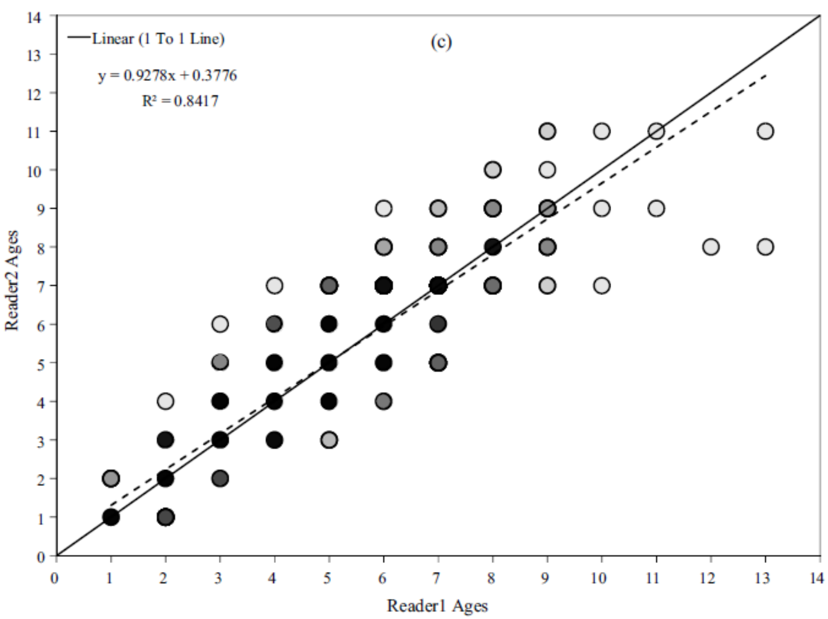
HEALY ET AL.




# Age reading precision




# FT-NIRS ageing precision