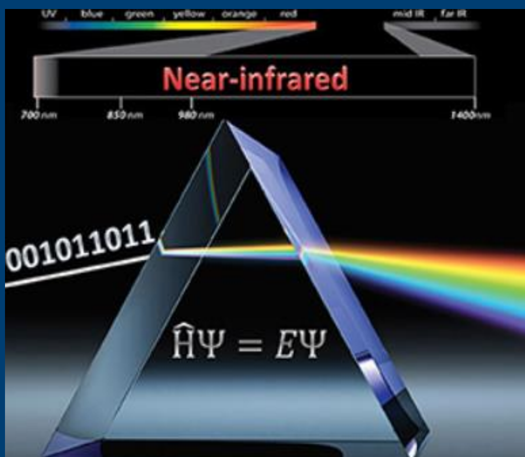
Science, Service, Stewardship



A revolutionary approach for improving age determination efficiency in fish using Fourier transform near infrared spectroscopy (FT-NIRS)

FT-NIRS 2023 Workshop 3-7 April

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

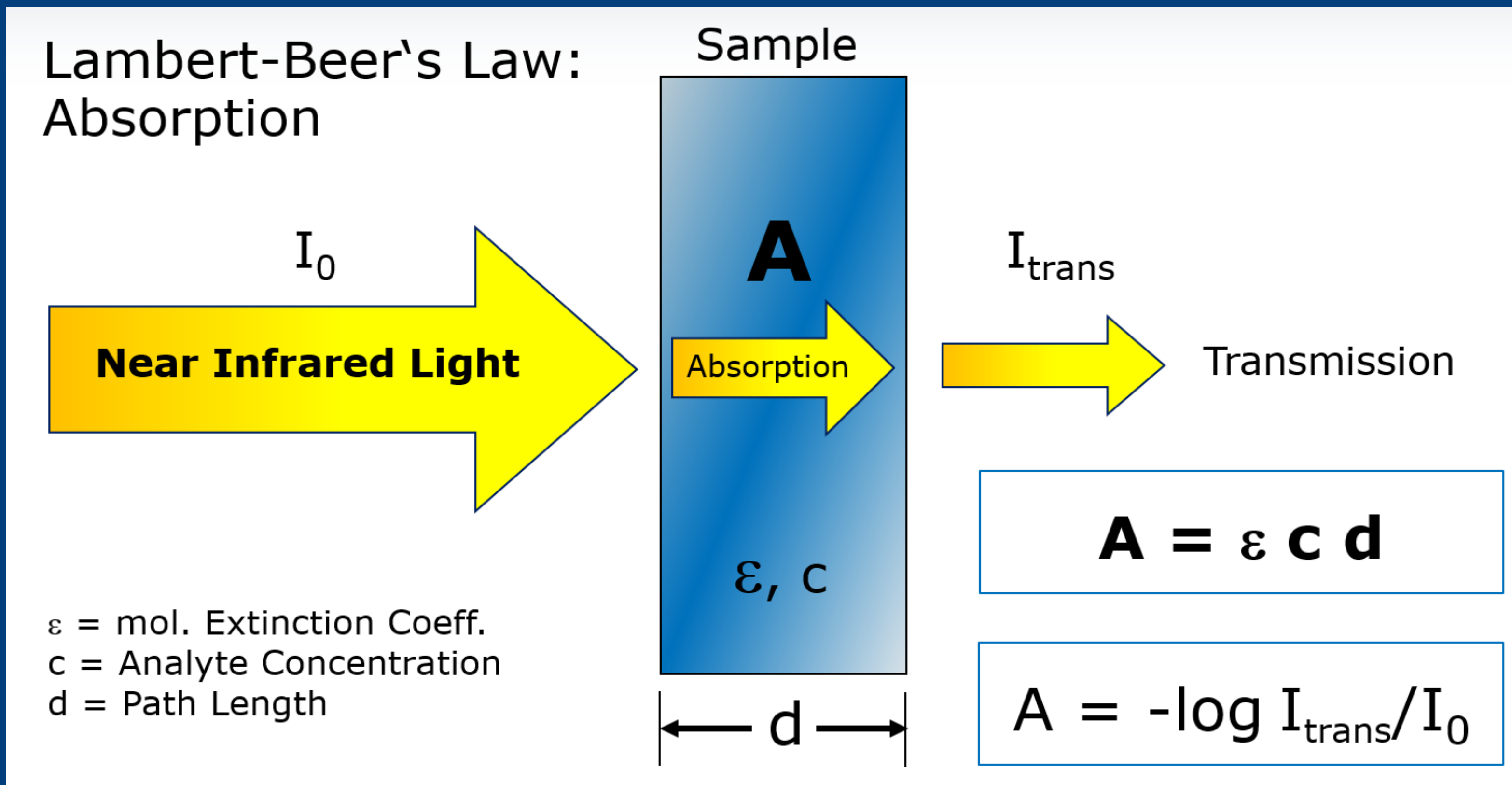
NOAA FISHERIES SERVICE

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

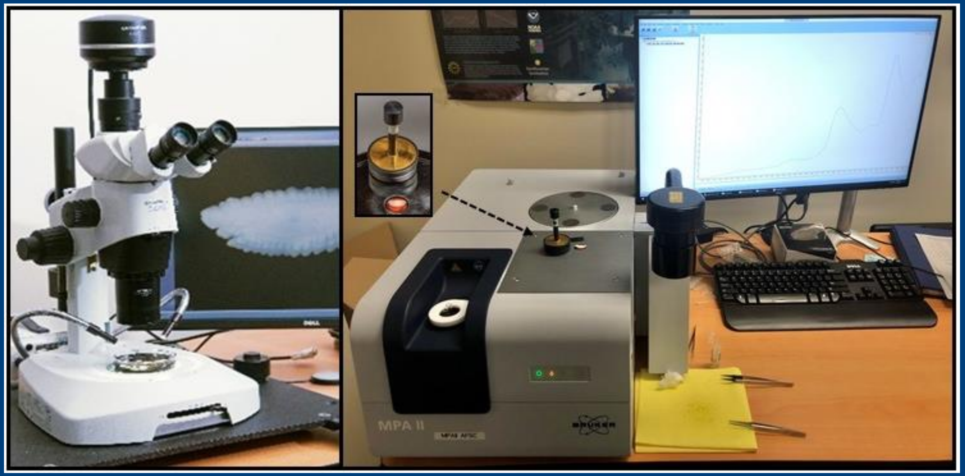
Where to begin?

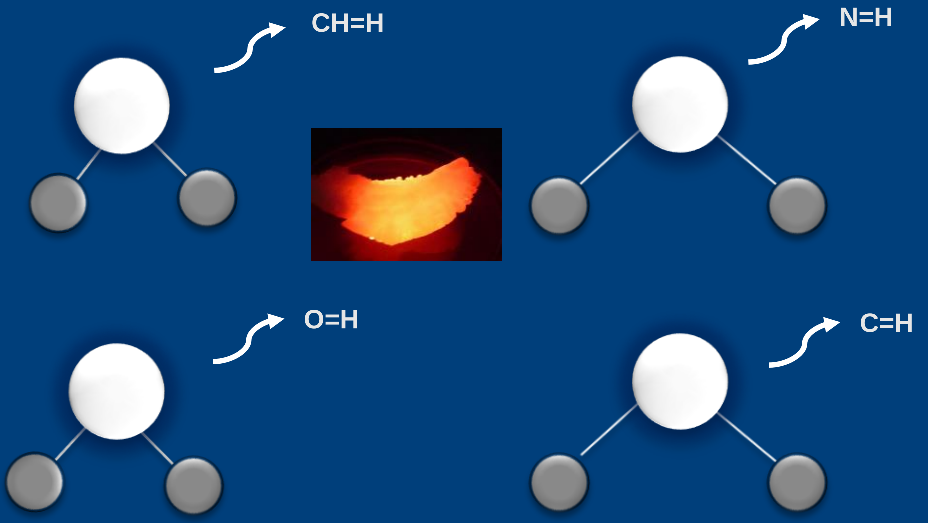
- Some background of FT-NIRS & age prediction
- How does FT-NIRS age estimation fit into the TMA process (maintain consistency in both TMA and FT-NIRS)
- Describe a (***hypothetical***) system of process control and data quality control from end-to-end (reporting examples)
- Emulate the future process: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

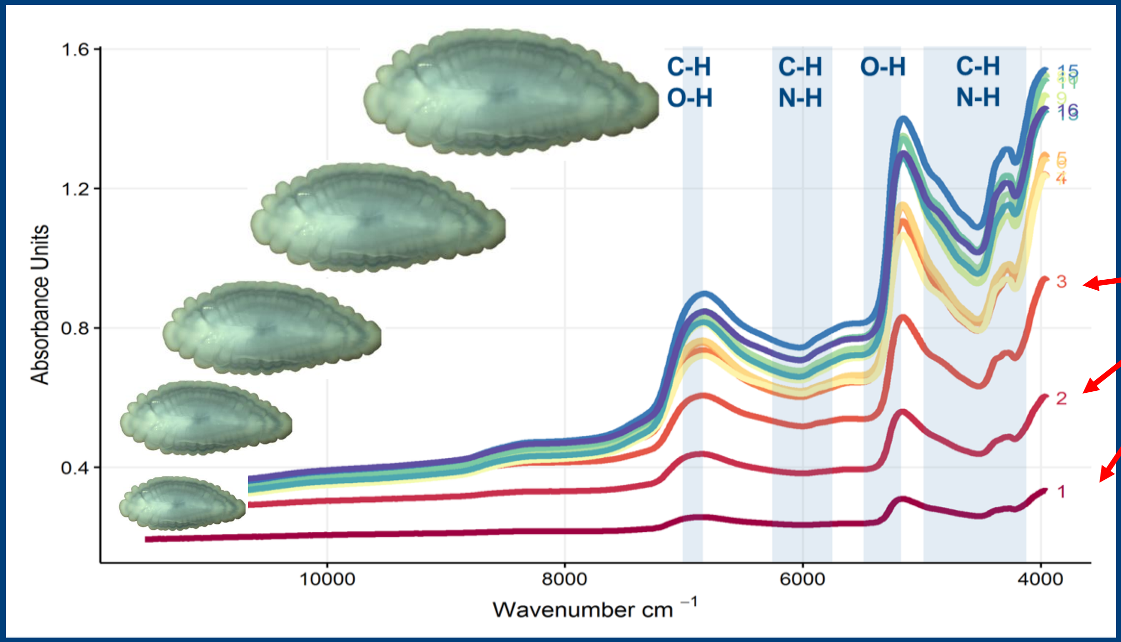
Interaction of electromagnetic radiation and matter



NIR Spectroscopy: measurement of intra-molecular vibrations



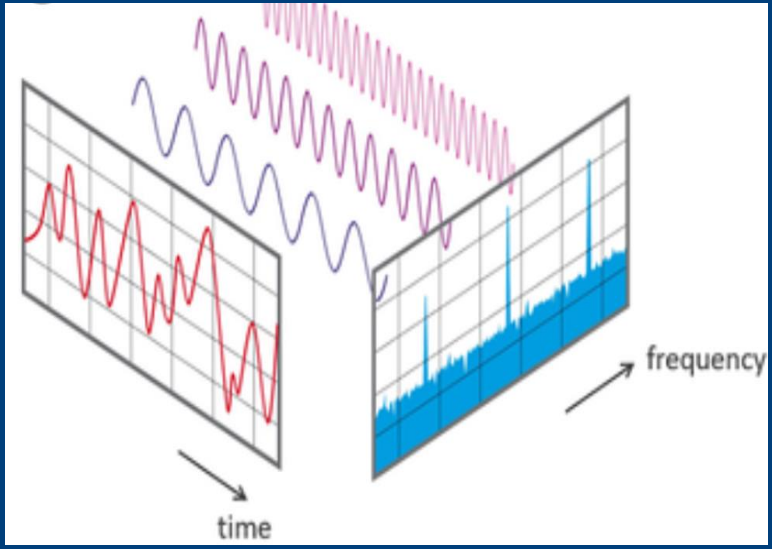




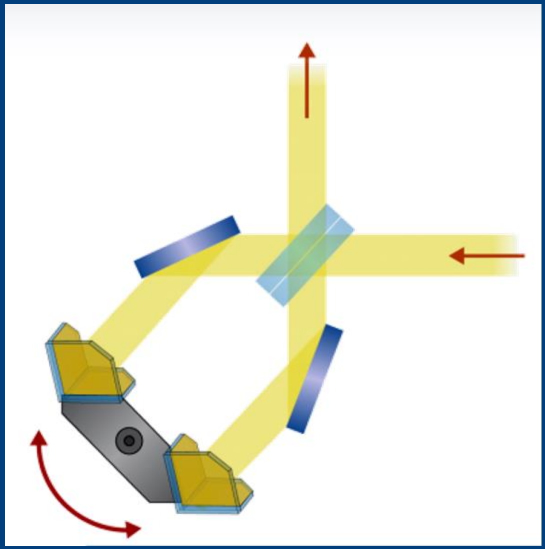
Different ages have different absorbance profiles

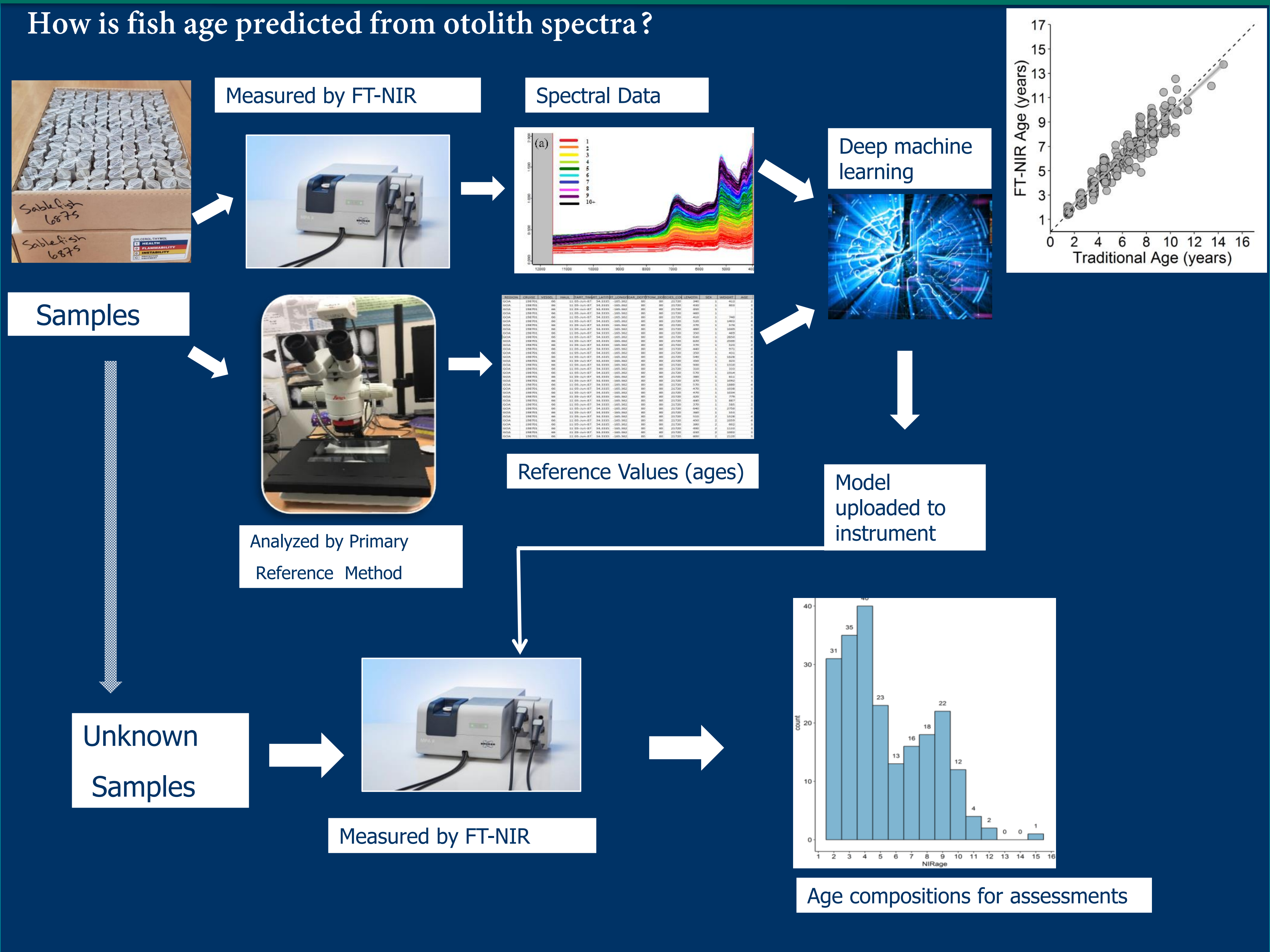
How are the molecular motions detected and measured?

Fourier transform near infrared spectrometer



Interferometer




How is fish age predicted from otolith spectra?
Sablefish 6875
Sablefish 6875
Measured by FT-NIR
Spectral Data
(a)
1
2
3
4
5
6
7
8
9
10+
Deep machine learning
FT-NIR Age (years)
Traditional Age (years)
Samples
Reference Values (ages)
Analyzed by Primary Reference Method
Model uploaded to instrument
Unknown Samples
Measured by FT-NIR
count
NIRage
Age compositions for assessments

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

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[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

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

esa

JNIRS

volume 27 / number 6 / December 2019
ISSN 0967-0335

JOURNAL OF NEAR INFRARED SPECTROSCOPY

Barium sulfate as a reference standard / Monitoring rice vinegar / Measuring strain in rocks

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

Journal of Near Infrared Spectroscopy
2022, Vol. 0(0) 1–10
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09670335221097005
journals.sagepub.com/home/jns
SAGE

Fourier transform near infrared spectroscopy as a tool to predict spawning status in Alaskan fishes with variable reproductive strategies

Todd TenBrink[1], Sandra Neidetcher[1], Morgan Arrington[2], Irina Benson[1], Christina Conrath[3] and Thomas Helser[1]

frontiers in Marine Science

ORIGINAL RESEARCH
published: 15 July 2021
doi: 10.3389/fmars.2021.690934

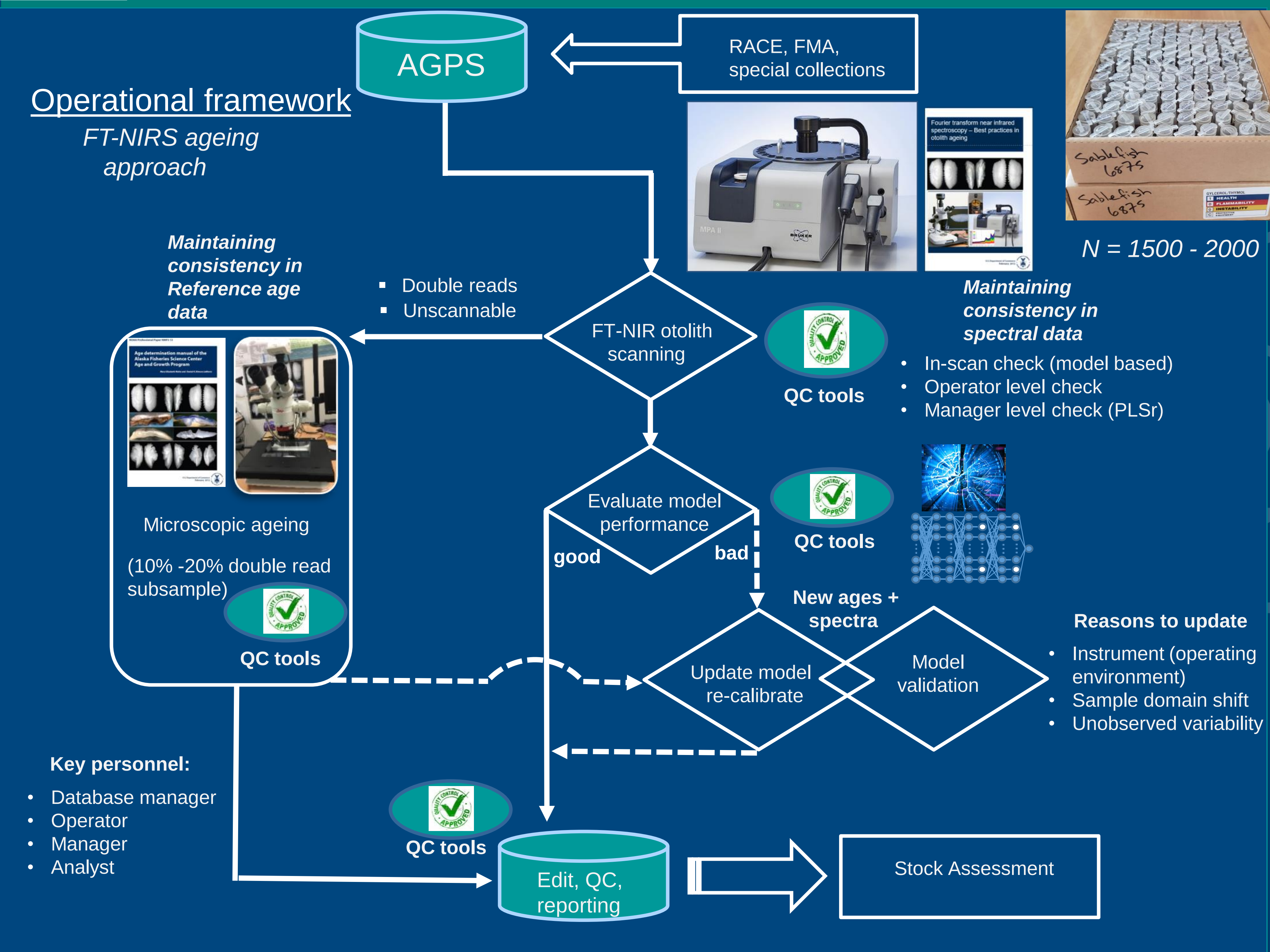
OPEN ACCESS

Edited by:
Carlo C. Lazado,
Norwegian Institute of Food, Fisheries and Aquaculture Research (Nofima), Norway

Reviewed by:

Rapid and Reliable Assessment of Fish Physiological Condition for Fisheries Research and Management Using Fourier Transform Near-Infrared Spectroscopy

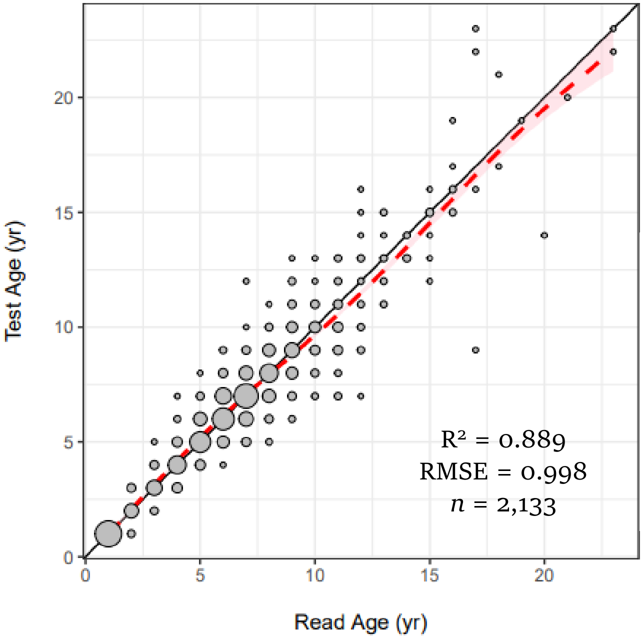
Esther D. Goldstein[1*], Thomas E. Helser[1], Johanna J. Vollenweider[2], Ashwin Sreenivasan[3] and Fletcher F. Sewall[2]


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

Pretend it is August of 2018.

- **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



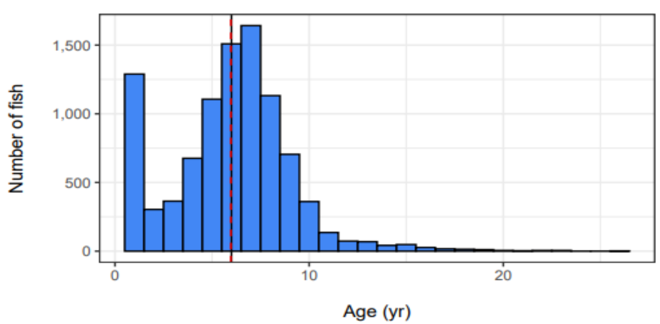
Precision statistic	Value
Percent agreement (PA)	66.9%
Average percent error (APE)	3.05%
Coefficient of variation (CV)	4.32%
Total number of fish in ageing collection	9603
Number of fish unaged	58
Number of fish in precision-testing sample	2133
Number of fish aged by two readers	2133
Percentage of fish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

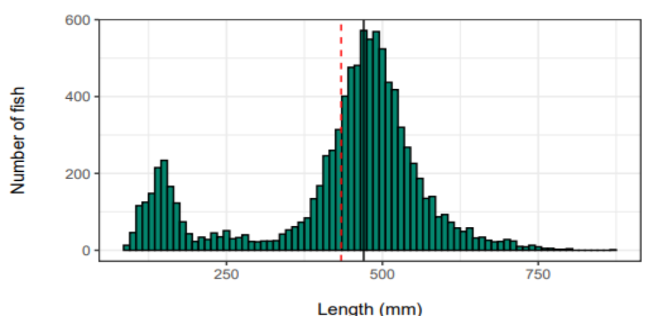
(a) Bias direction

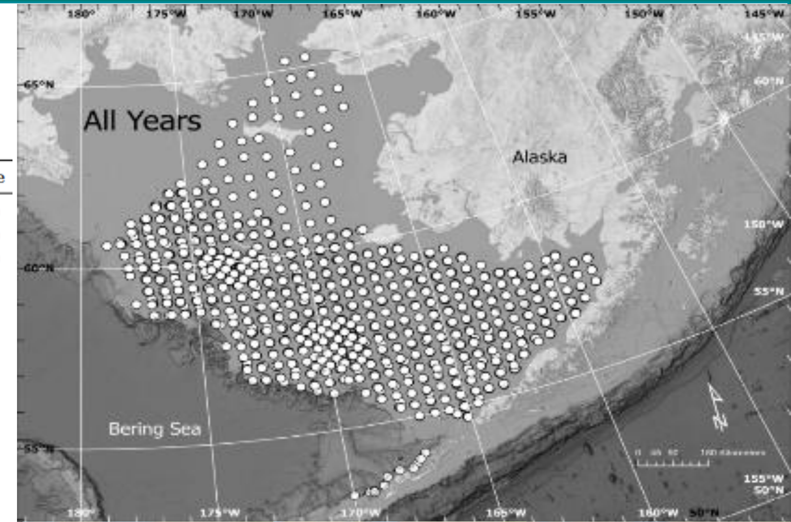
minus bias	371 otoliths	17.4%
plus bias	335 otoliths	15.7%
complete agreement	1427 otoliths	66.9%

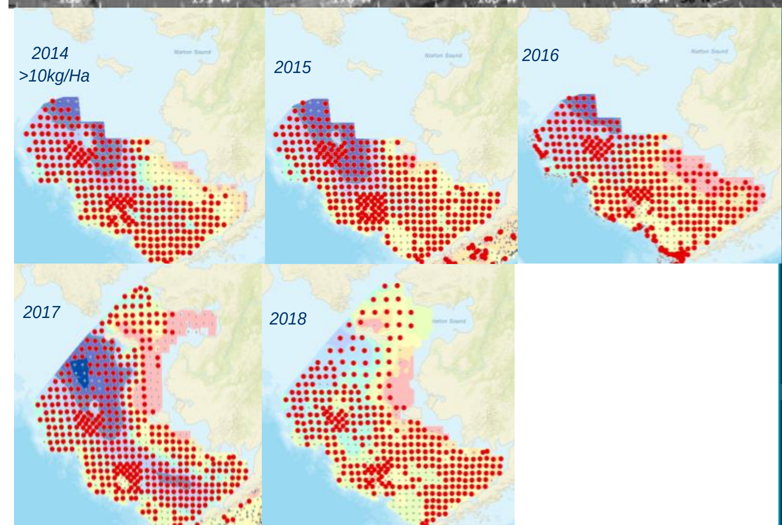
(b) Tests of symmetry

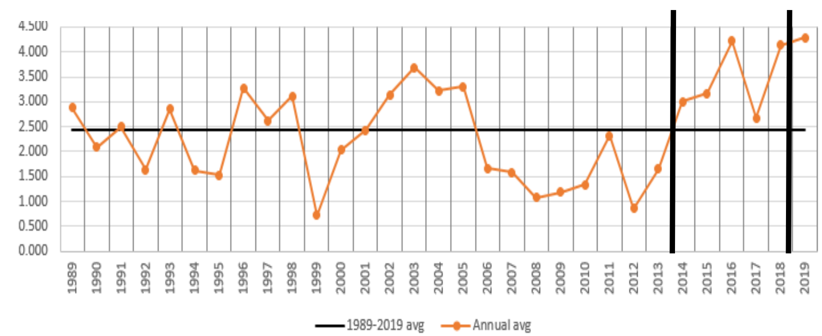
Test name	df	Test statistic	*p*
Bowker's	48	55.72	0.21
Evans-Hoenig	7	4.79	0.69





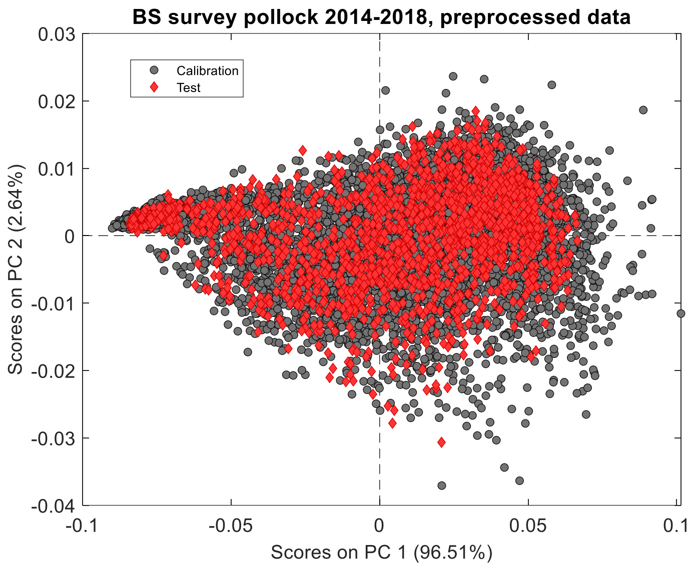




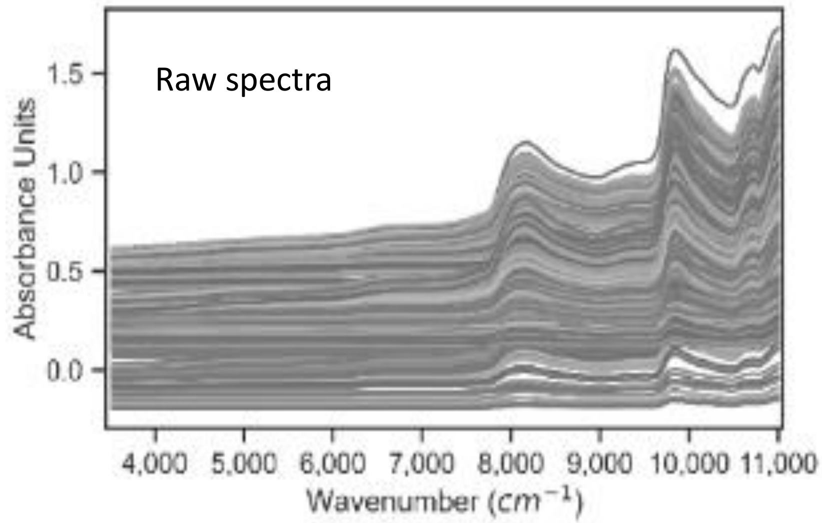


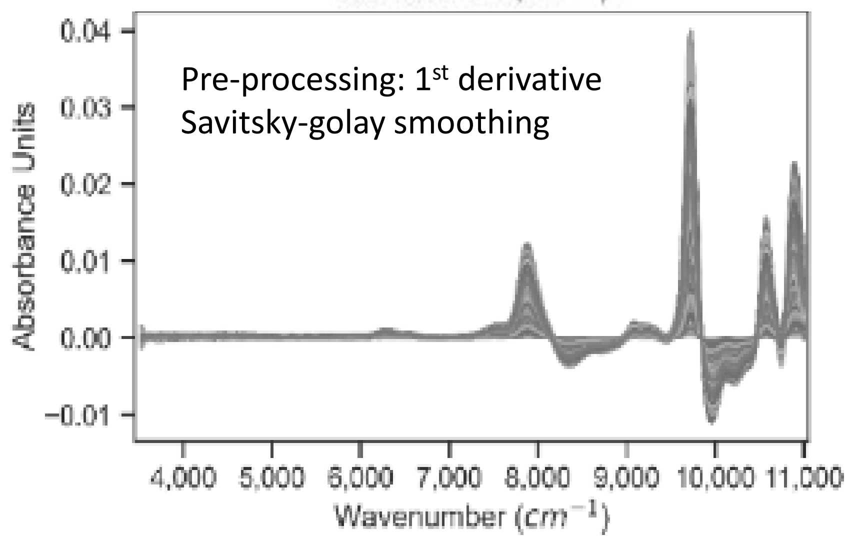
EBS walleye pollock base model (2014-2018)
Otolith spectra

- n = 8,617
- Onion algorithm to split train & test (applied to each year)



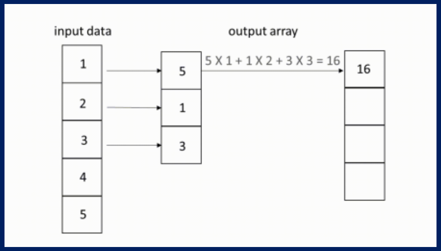
Features	Subset	Values
Mean fish age* (standard deviation)	train	5.72 (3.3)
	test	5.82 (3.14)
Fish age range*	train	1 – 23
	test	1 – 23
Fish length range	train	90 – 870
	test	100 – 800
Latitude range	train	53.13822 – 65.25117
	test	53.15808 – 65.23351
Gear depth range	train	18 – 695.79
	test	18 – 640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





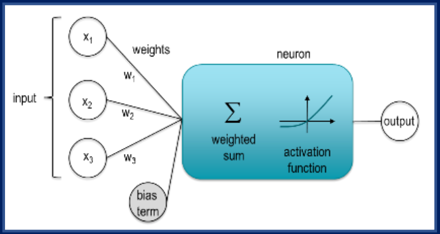
Convolutional Neural Network

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

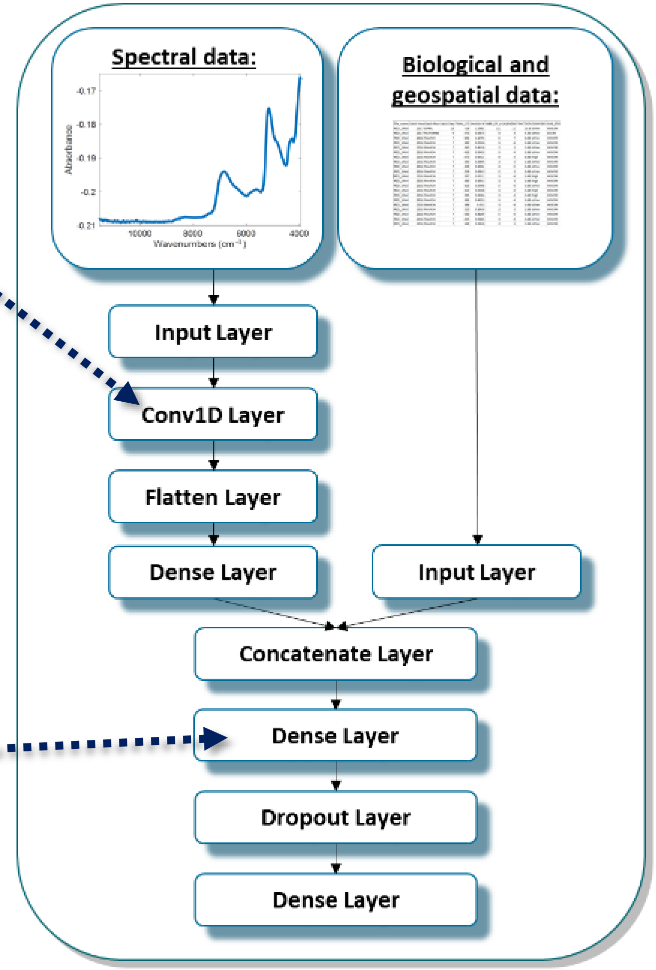


Deep learning networks will have multiple kernels and will produce multiple output arrays.

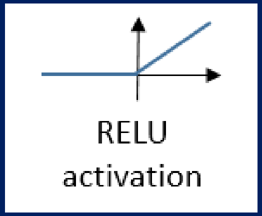
Neurons are core processing units of the network.



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

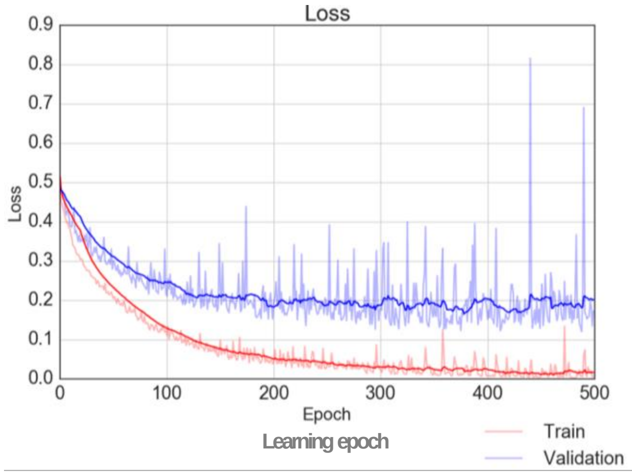


Non-linear activation functions introduce non-linear properties into network.



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 Science Publishing | Canadian Journal of Fisheries and Aquatic Sciences

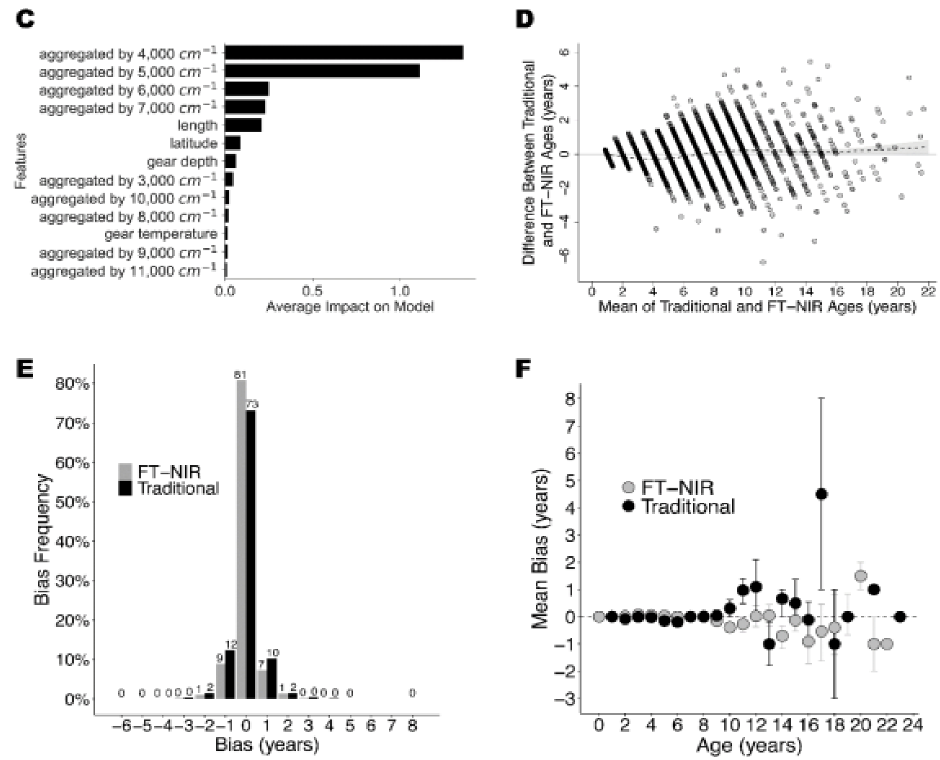
The future of fish age estimation: deep machine learning coupled with Fourier transform near infrared spectroscopy of otoliths.

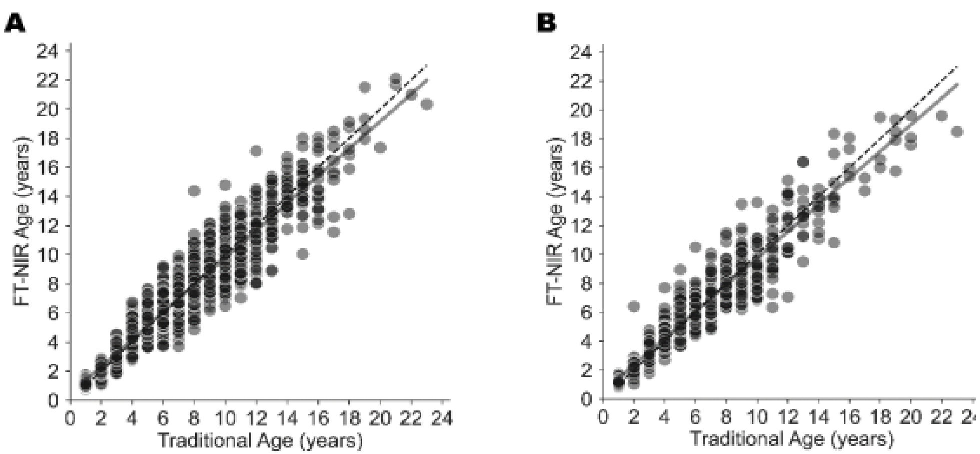
Benson, Irina; NOAA Fisheries Alaska Fisheries Science Center, Resource Ecology and Fisheries Management Division
Helser, Thomas; NOAA Fisheries Alaska Fisheries Science Center, Resource Ecology and Fisheries Management Division
Marchetti, Giovanni; Google LLC
Barnett, Beverly; NOAA Fisheries Southeast Fisheries Science Center, Fisheries Assessment, Technology, and Engineering Support Division, Biology and Life History Branch, Panama City Facility

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

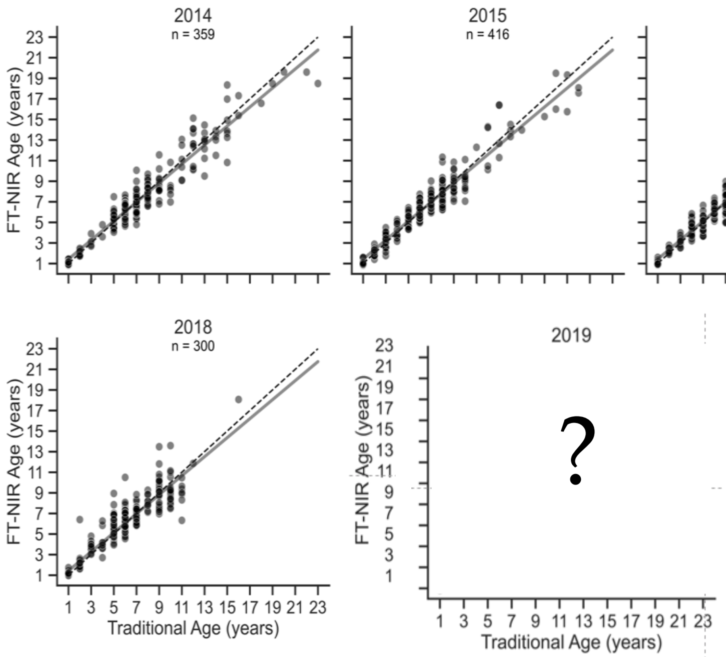
Model	Number of otoliths			R^2		RMSE	
	Train	Test	Outliers*	Train	Test	Train	Test
CNN	6866	1751	12	**0.93**	**0.92**	**0.83**	**0.91**
PLS with all spectral wavenumbers				0.89	0.87	0.99	1.14
PLS with selected spectral wavenumbers				0.90	0.87	0.97	1.12

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

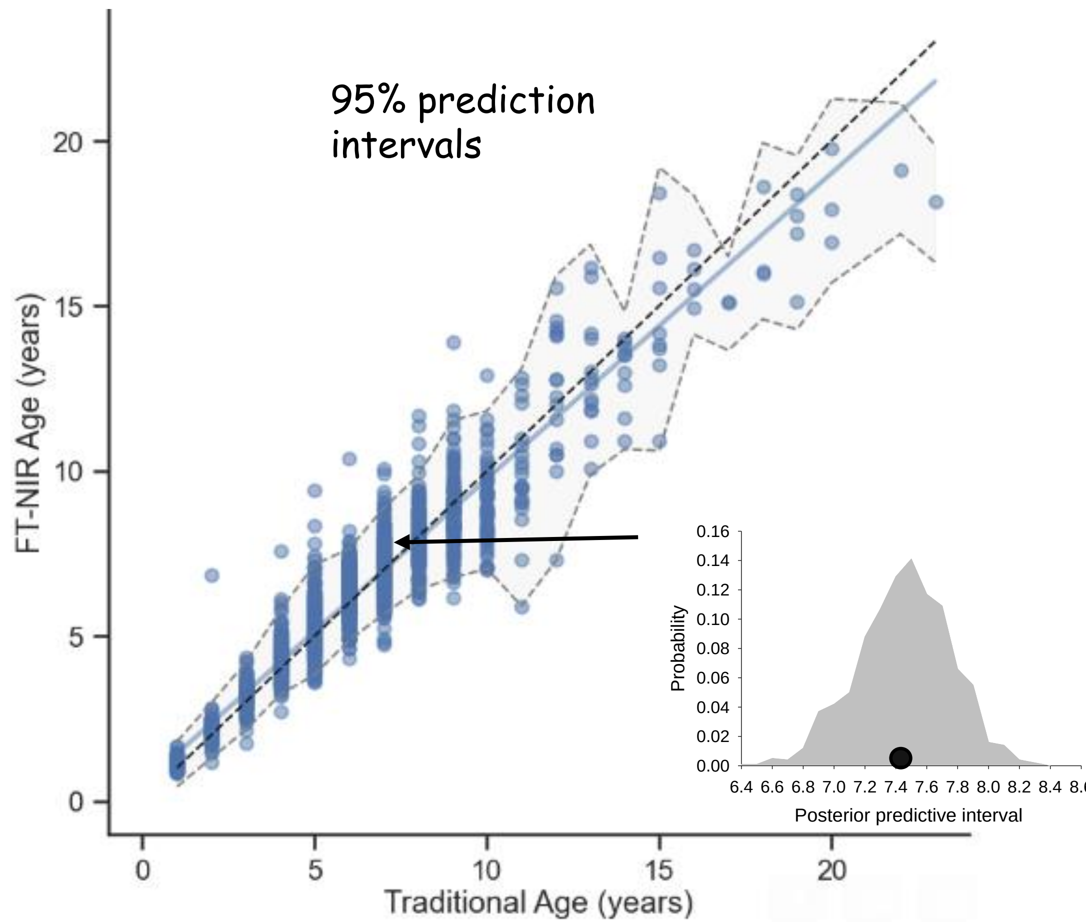




R^2 = 0.92 $RMSE_p$= 0.91 years



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

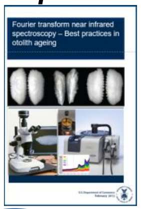
Nature Reviews Physics **4**, 573–577 (2022) | Cite this article

4138 Accesses | **23** Altmetric | Metrics

PART I: Collection of pollock otolith NIR spectra (*Still in 2019, but jump to the future to 2021*)

Consistency in otolith spectra

FT-NIR otolith scanning



QC tools

QA/QC standardized reports ex. 2019 survey walleye pollock

Analysis by session	Total scans	Unscannable	Broken	Crystallized	Other issue	Outliers	Extreme	Unusual
NIR_162201901201A	138	2	2	3	5	2	3	4
NIR_162201901201B	261	11	1	10	16	2	12	6
NIR_162201901201C	140	0	4	4	5	1	9	4
NIR_162201901201D	140	0	0	6	9	1	12	6
NIR_162201901201E	85	0	0	1	2	0	7	2
NIR_94201901201A	124	7	6	6	4	0	10	6
NIR_94201901201B	236	44	13	10	30	1	24	4
NIR_94201901201C	138	2	0	7	7	1	11	4
NIR_94201901201D	138	2	2	6	7	1	11	2
NIR_94201901201E	78	0	0	0	3	0	4	4
Separate analysis on combined 2019 scan sessions	1478	68	28	53	88	5	82	42

Outliers – Rescan or traditional ageing

Extreme – Pass QA/QC if there are no issues based on metadata

Unusual –
- Confirm outliers
- Possibly add data to calibration model
  - 2019 warm year
  - Many small fish

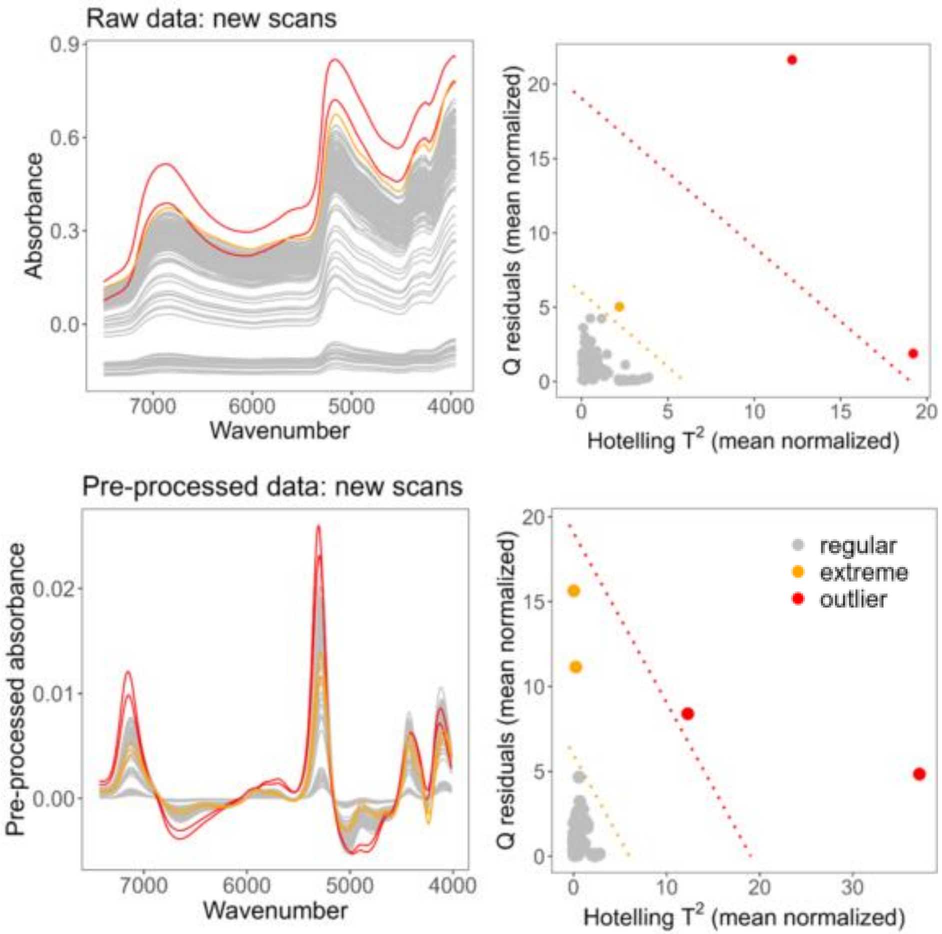
Include time-flow stats to monitor utilization and efficiency:

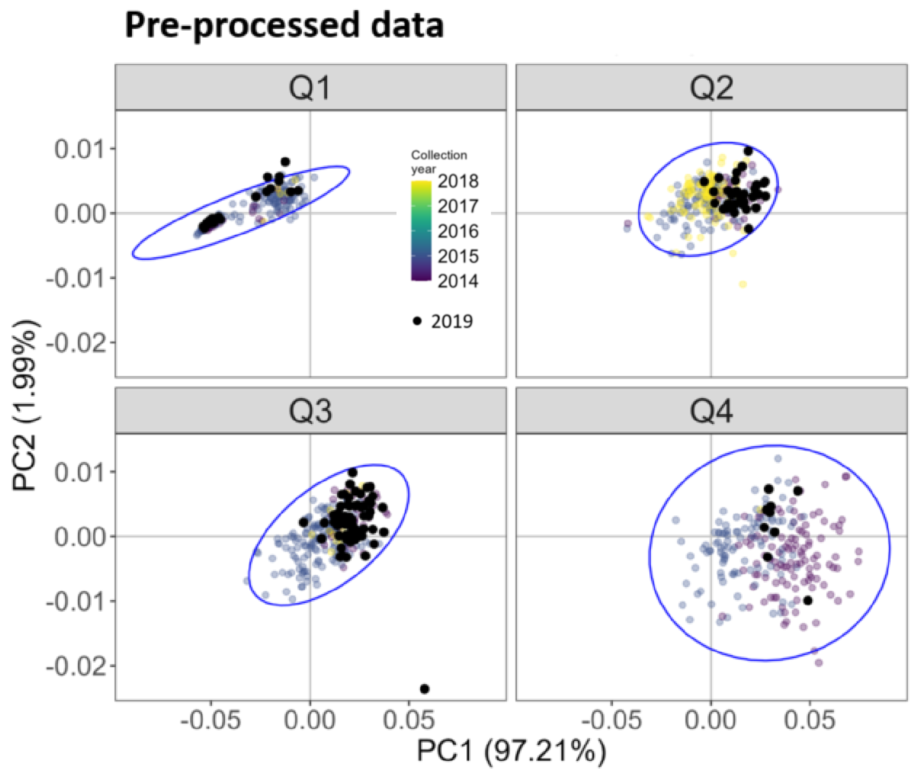
- 10 operators
- 1.66 – 2.43 hrs/session
- 30 – 50 scans/person/hour (as high as 67 scans/hour)

Future scan sessions – standardized at ~ 2 boxes (200+ otos)

Higher efficiency may be realized (w/o weighing otoliths)

QA/QC standardized reports
ex. 2019 survey walleye Pollock Session NIR_162201901201A





Part 1: Extremes & Outliers
PCA by scan session

→ Extreme: determine cause (small otolith, broken, etc.)
Significance level = 0.05

→ Outliers: rescan or age traditionally.
Significance level = 0.01

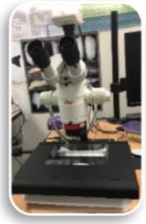
Part 2: Unusual spectra
New scans vs. past scans

- PCAs with all data
- 99% data ellipse by length quartile

→ outliers & specimens that expand data domain

PART II: Processing 2019 & 2021 age data
(only ageing 20% of entire collection)

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

QC tools

2019

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

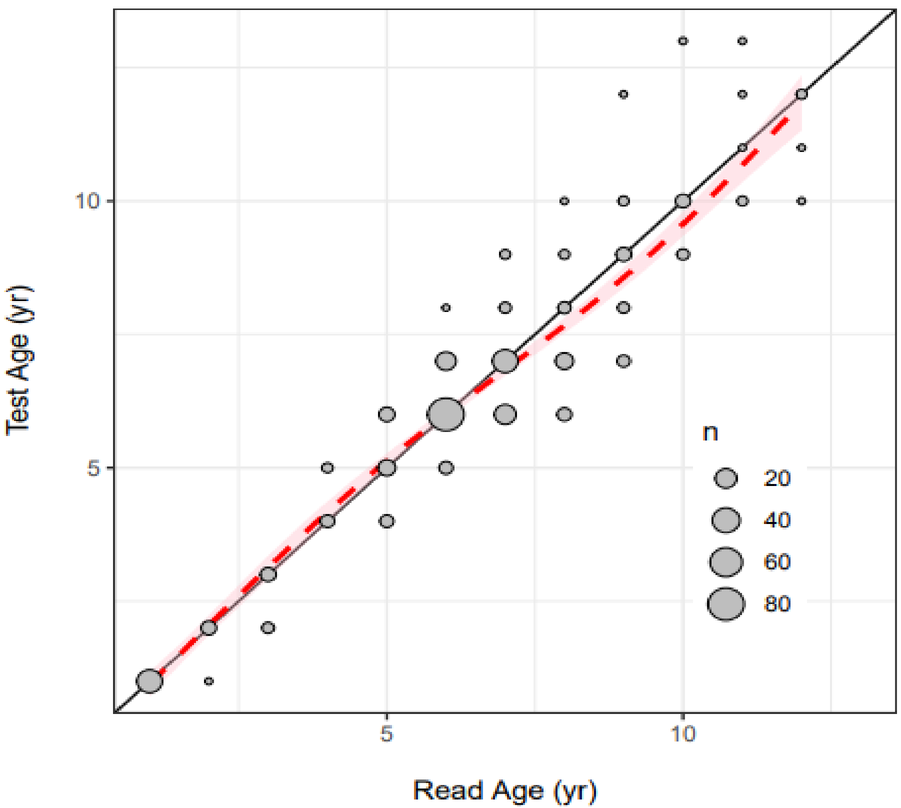
Precision statistic	Value
Percent agreement (PA)	64.61%
Average percent error (APE)	3.2%
Coefficient of variation (CV)	4.53%
Total number of fish in ageing collection	1552
Number of fish unaged	13
Number of fish in precision-testing sample	308
Number of fish aged by two readers	308
Percentage of fish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

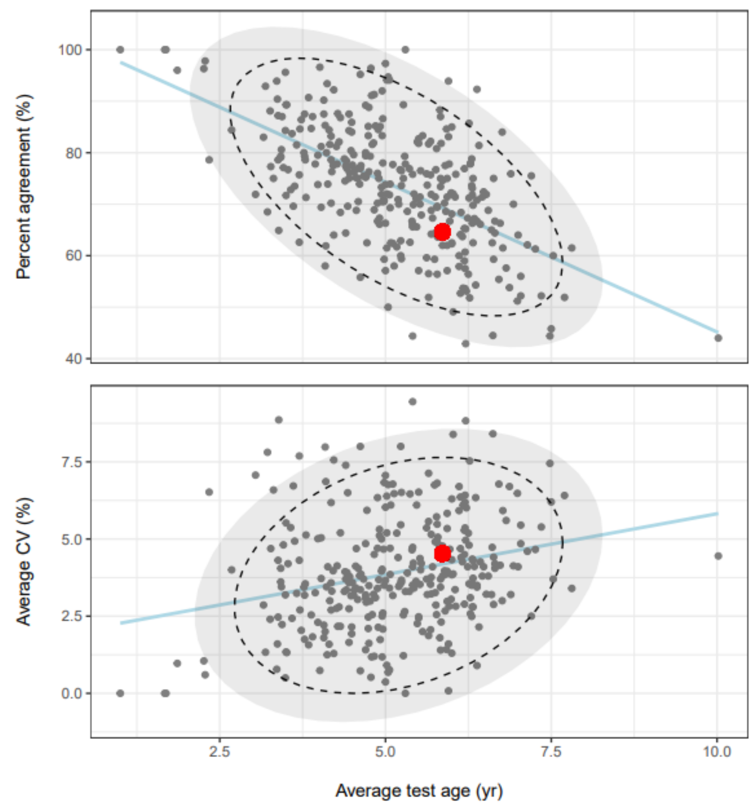
(a) Bias direction

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

(b) Tests of symmetry

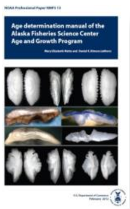
Test name	df	Test statistic	p
Bowker's	17	23.44	0.14
Evans-Hoenig	3	9.1	0.03

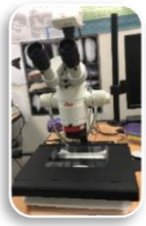




2021

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





QC tools

Table 1: Precision statistics for the Walleye Pollock 2021 Bering Sea Survey collection.

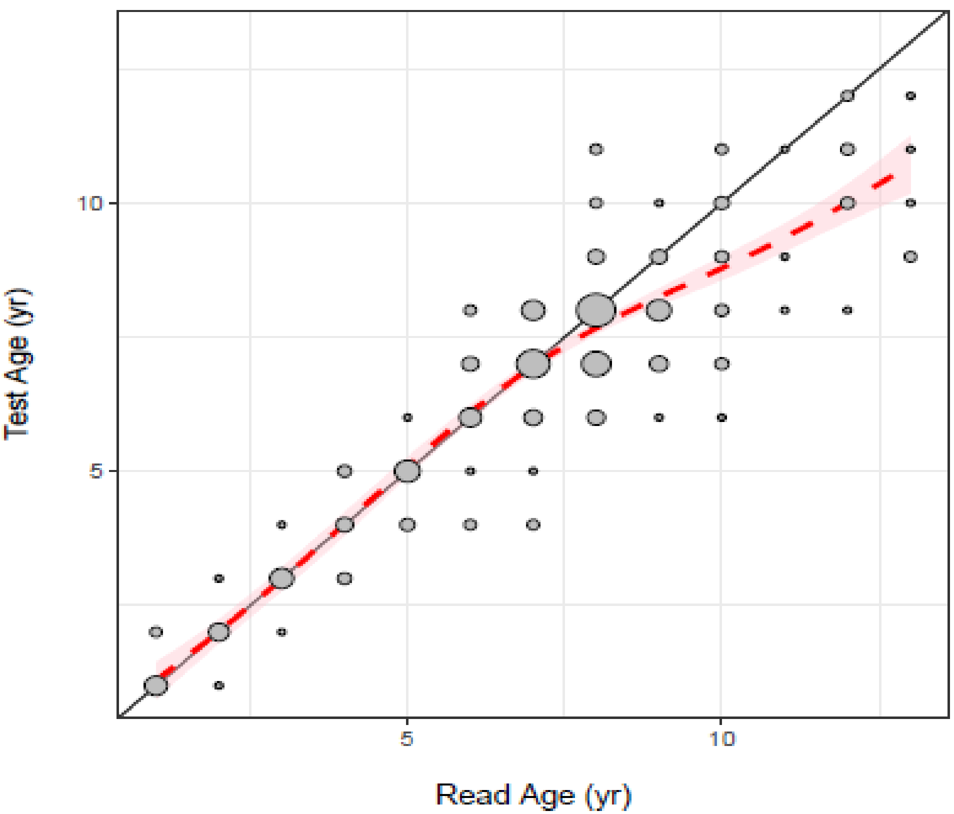
Precision statistic	Value
Percent agreement (PA)	56.27%
Average percent error (APE)	4.34%
Coefficient of variation (CV)	6.14%
Total number of fish in ageing collection	1535
Number of fish unaged	7
Number of fish in precision-testing sample	359
Number of fish aged by two readers	359
Percentage of fish with paired age readings	23%
Average read age (paired reads only)	6.89
Average test age (paired reads only)	6.57

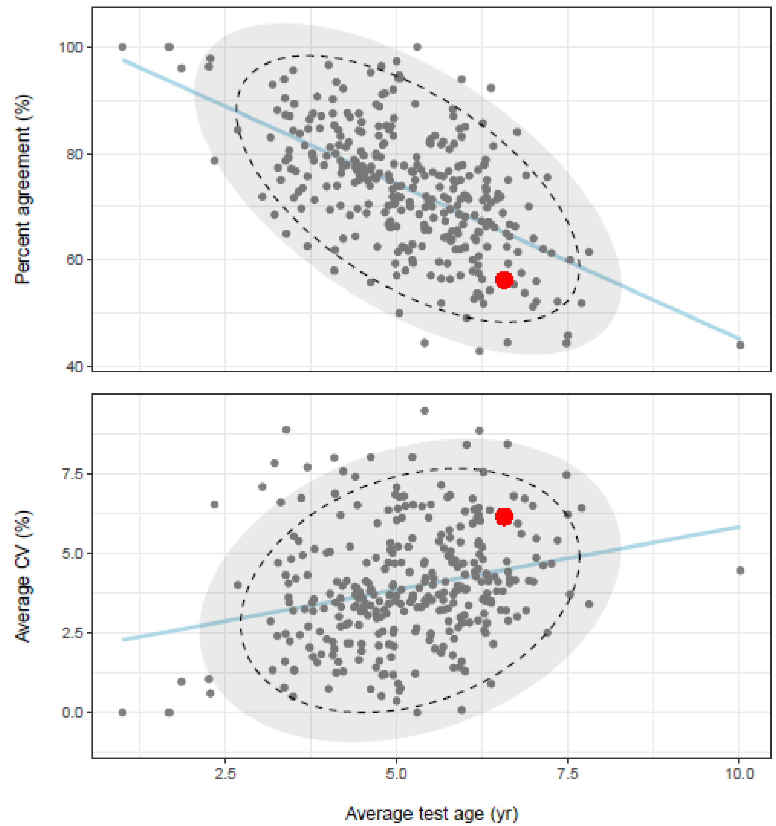
(a) Bias direction

minus bias	45 otoliths	12.5%
plus bias	112 otoliths	31.2%
complete agreement	202 otoliths	56.3%

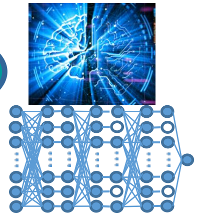
(b) Tests of symmetry

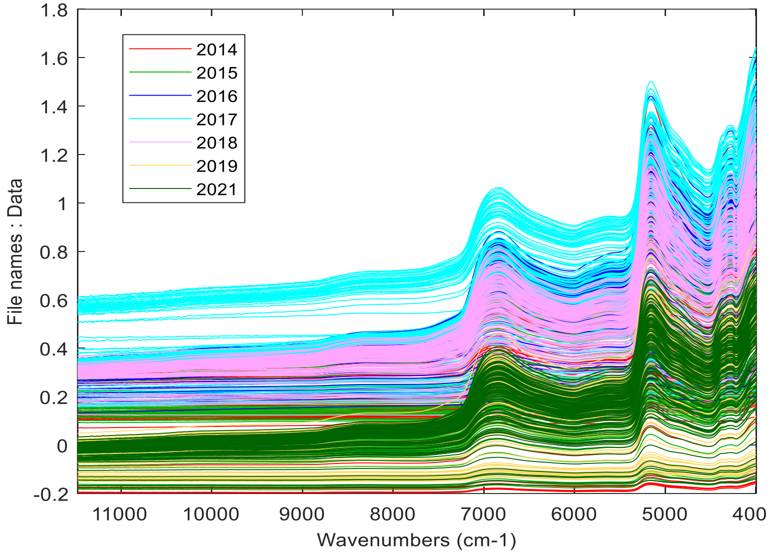
Test name	df	Test statistic	*p*
Bowker's	28	50.09	0.01
Evans-Hoenig	4	34.57	0



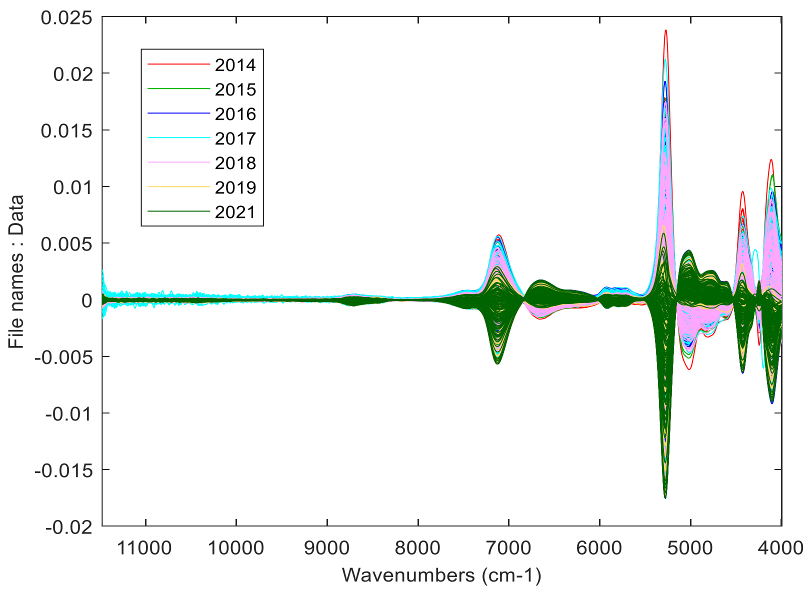


QC tools

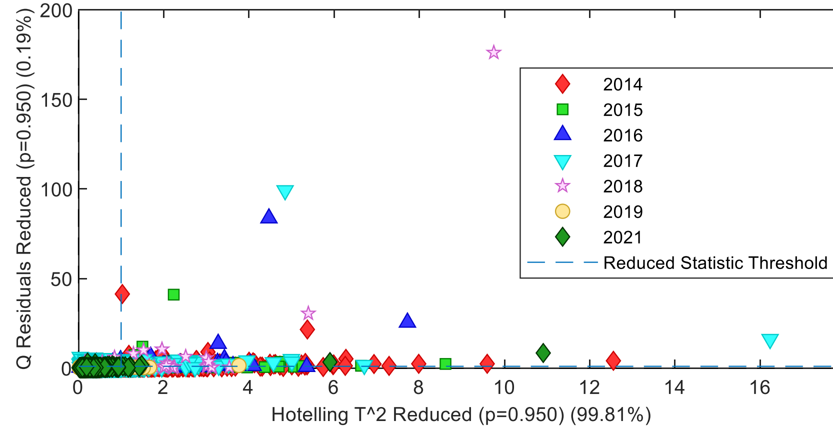




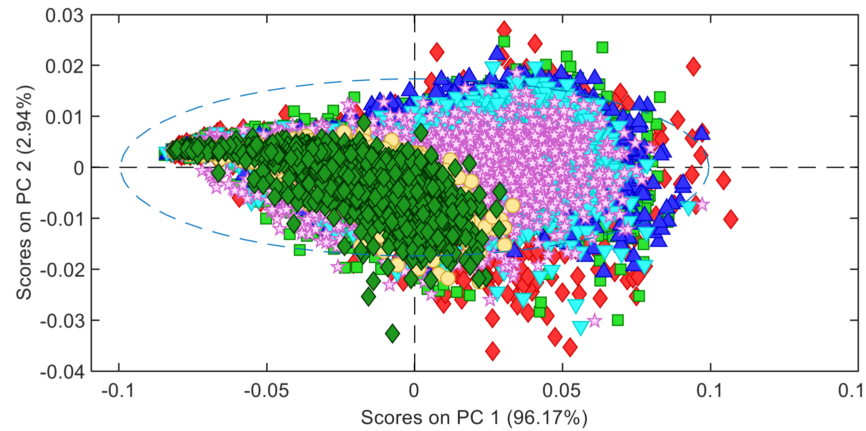
File names : Data
Wavenumbers (cm-1)
2014
2015
2016
2017
2018
2019
2021



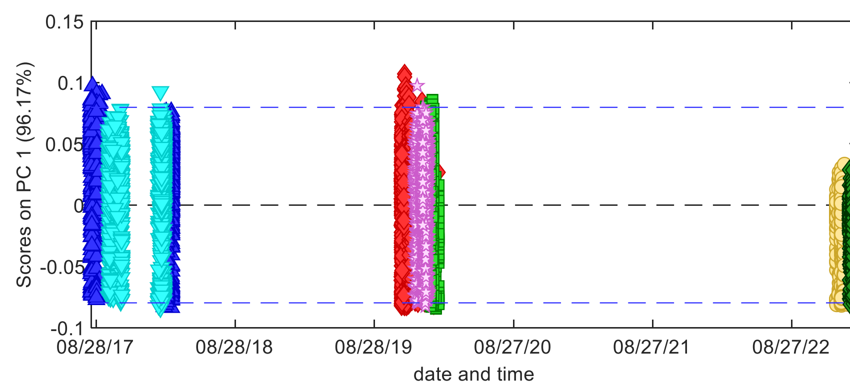
File names : Data
Wavenumbers (cm-1)
2014
2015
2016
2017
2018
2019
2021



Q Residuals Reduced (p=0.950) (0.19%)
Hotelling T^2 Reduced (p=0.950) (99.81%)
2014
2015
2016
2017
2018
2019
2021
Reduced Statistic Threshold

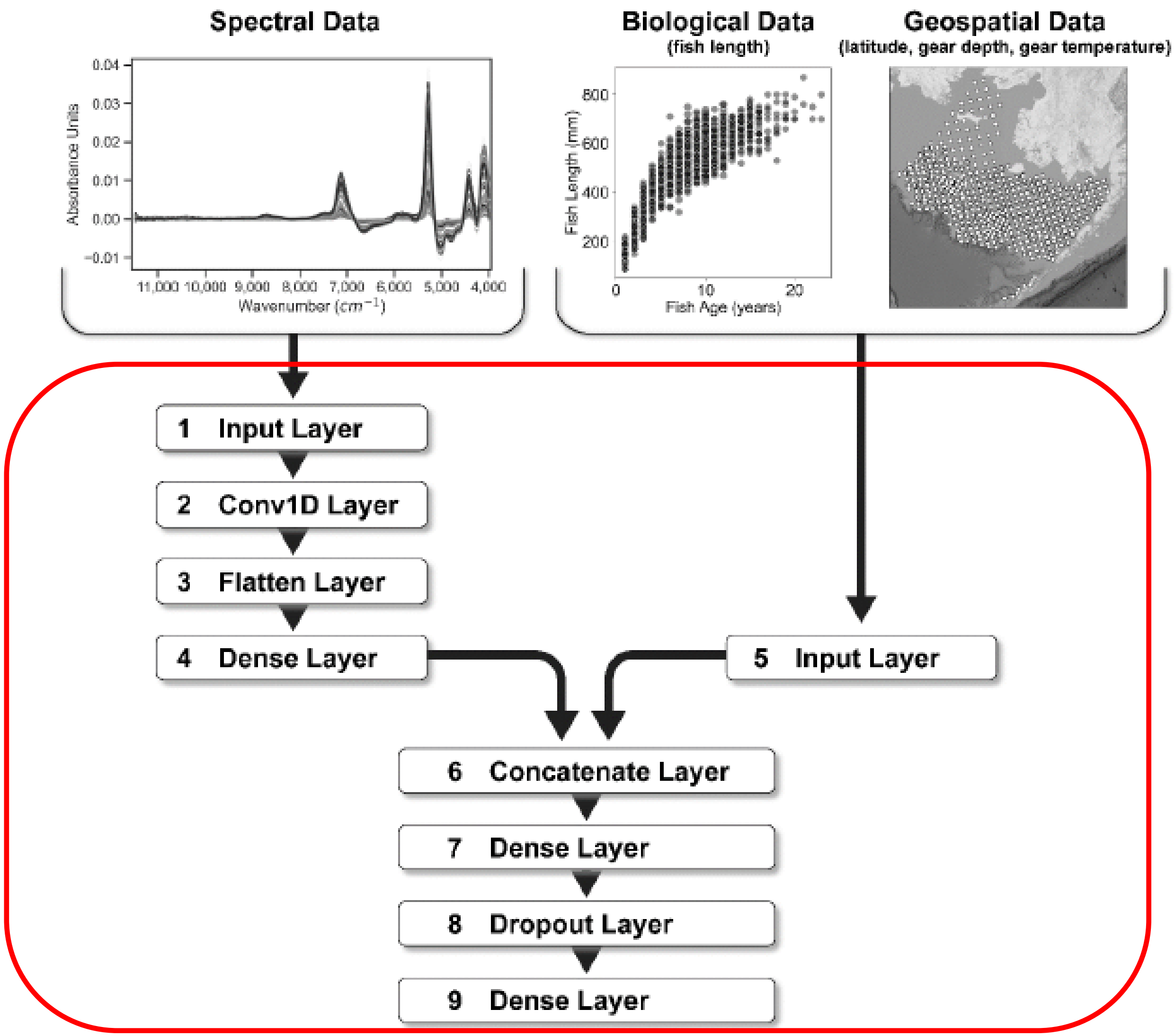


Scores on PC 2 (2.94%)
Scores on PC 1 (96.17%)

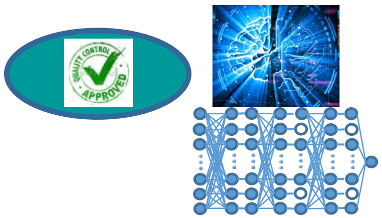


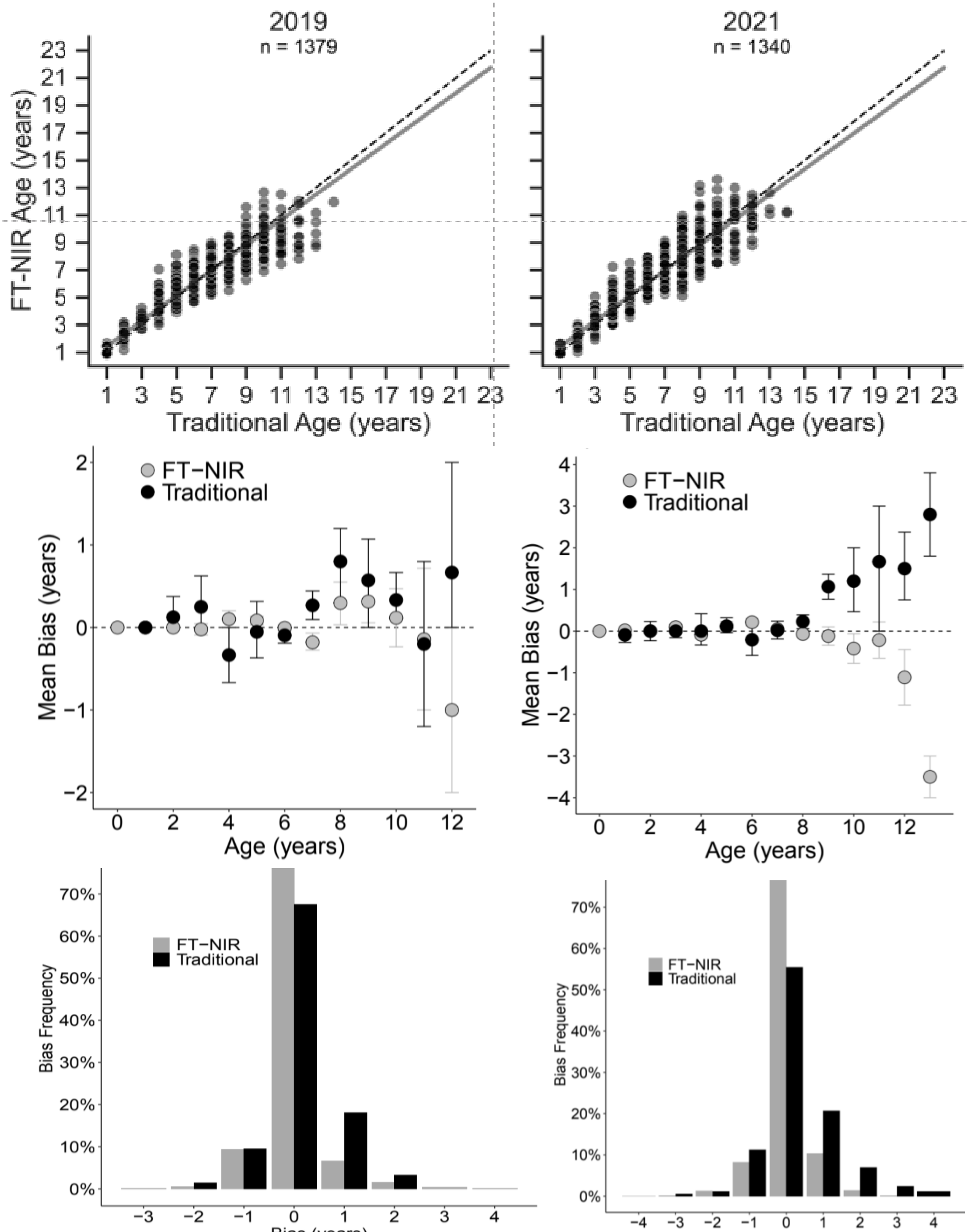
Scores on PC 1 (96.17%)
08/28/17
08/28/18
08/28/19
08/27/20
08/27/21
08/27/22
date and time

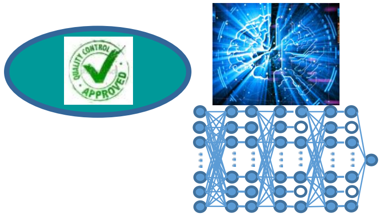
Predicting 2019 & 2021 (*ages, spectra & metadata*)



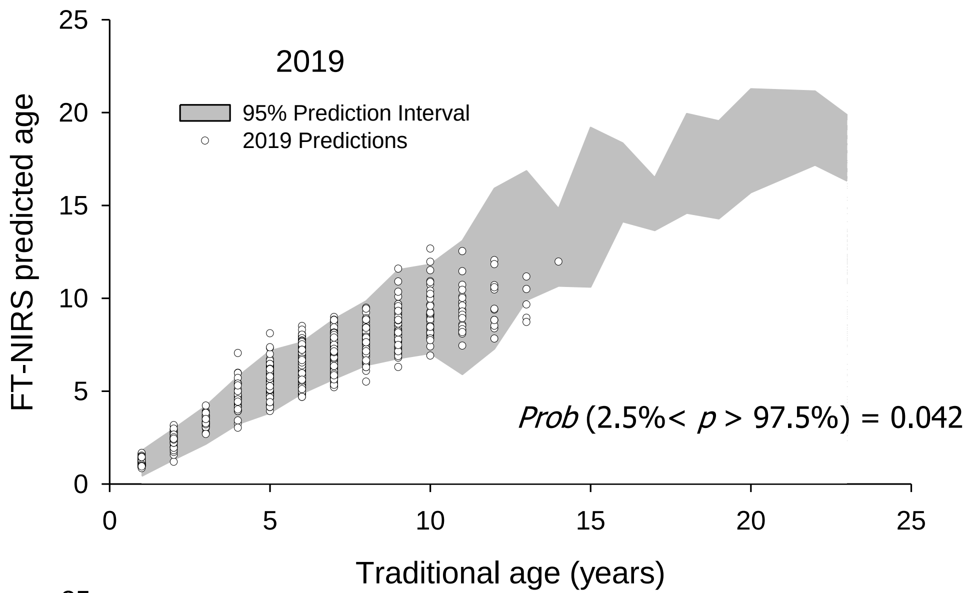
2014-2018 CNN base model

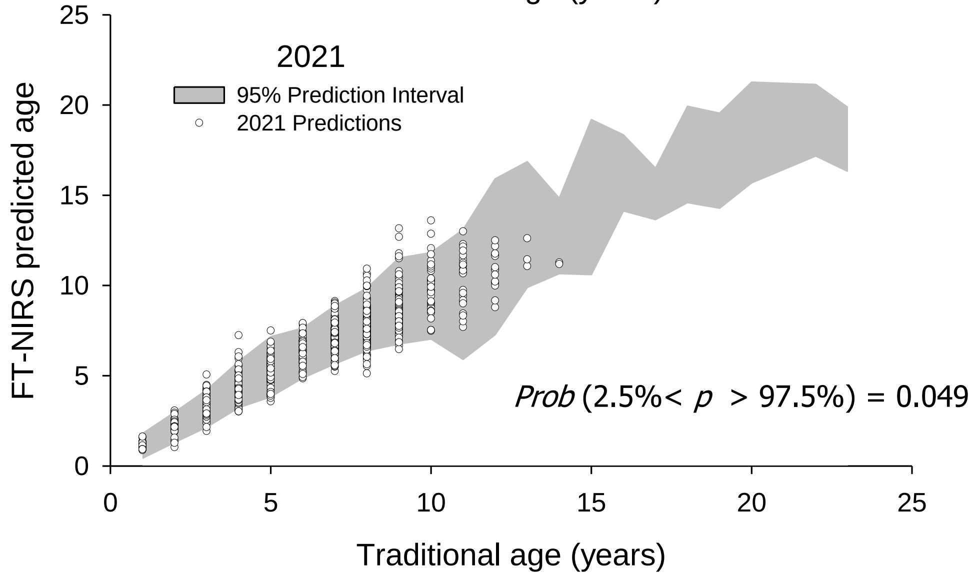

QUALITY CONTROL
APPROVED


2019
n = 1379
2021
n = 1340
FT-NIR Age (years)
Traditional Age (years)
Mean Bias (years)
Age (years)
FT−NIR
Traditional
Bias Frequency
Bias (years)

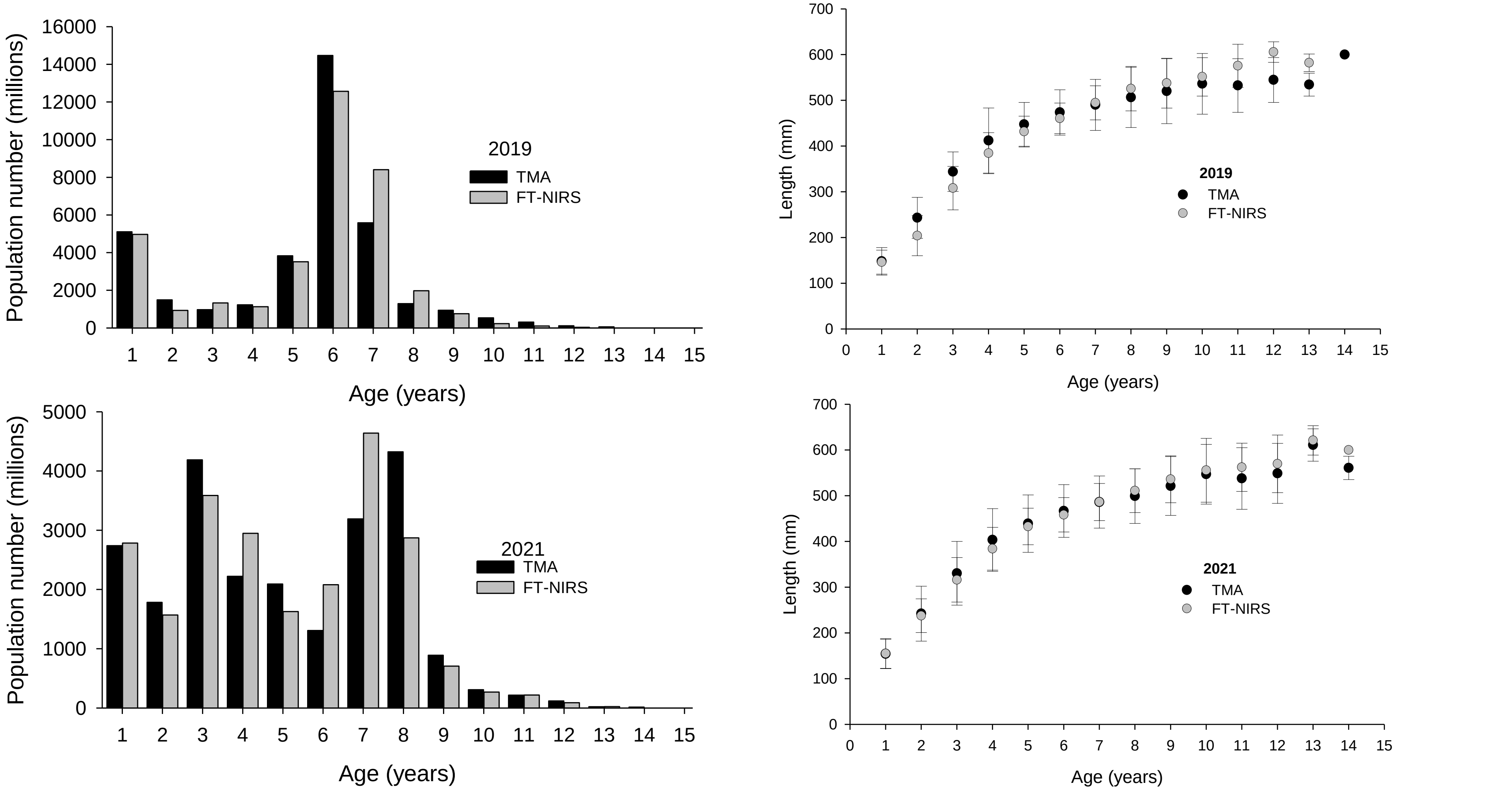


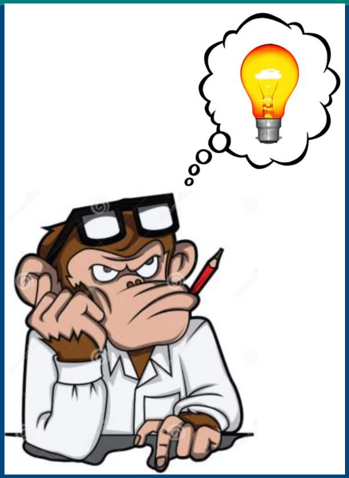
Assessing CNN model performance of future predictions





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?

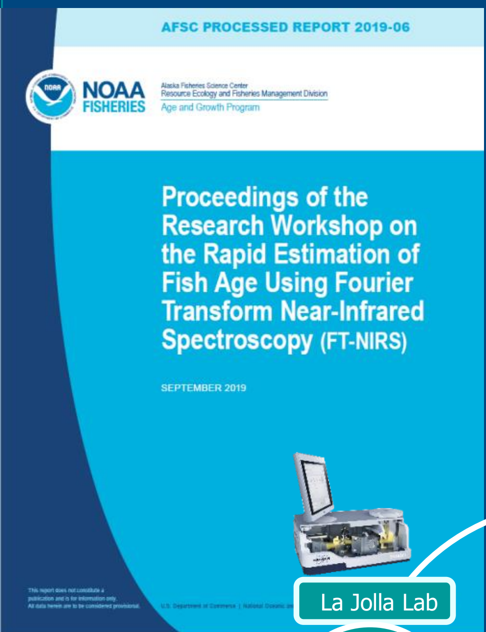
Strategic Initiative work flow

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 dept
- •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

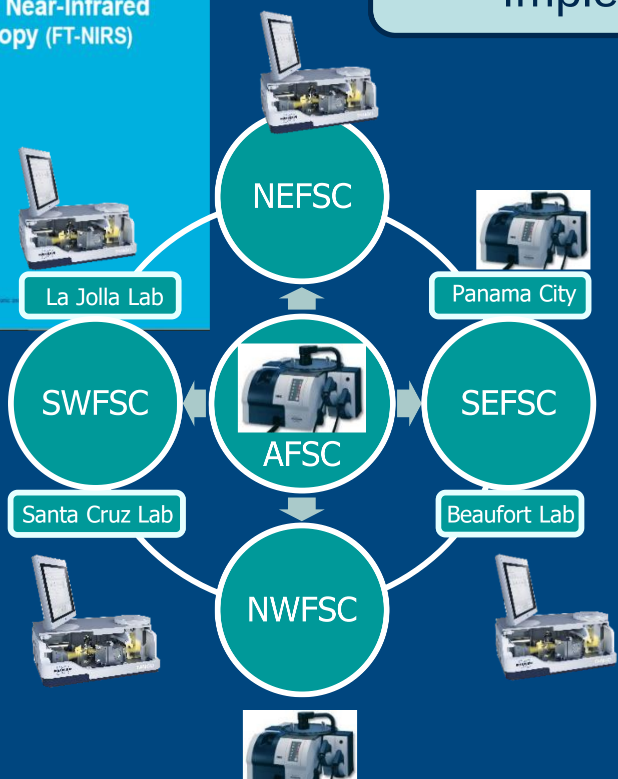
Proceedings of the Research Workshop on the Rapid Estimation of Fish Age Using Fourier Transform Near-Infrared Spectroscopy (FT-NIRS)

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



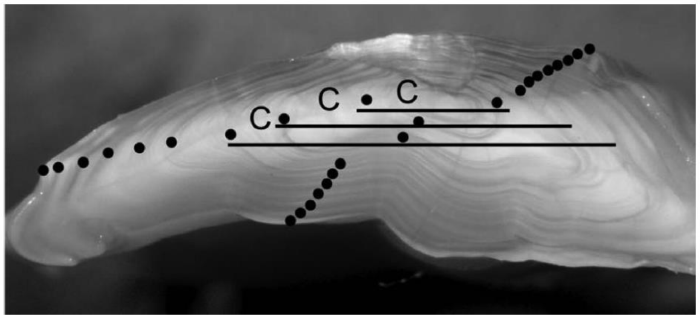
Stock Assessment Integration

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

Deployment of technology

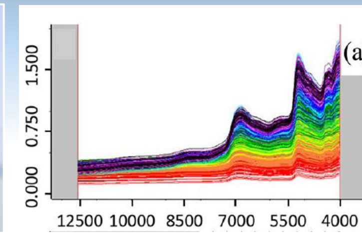
- •Integrating technology into current production setting (species-specific)

Pacific cod otolith (transverse section)

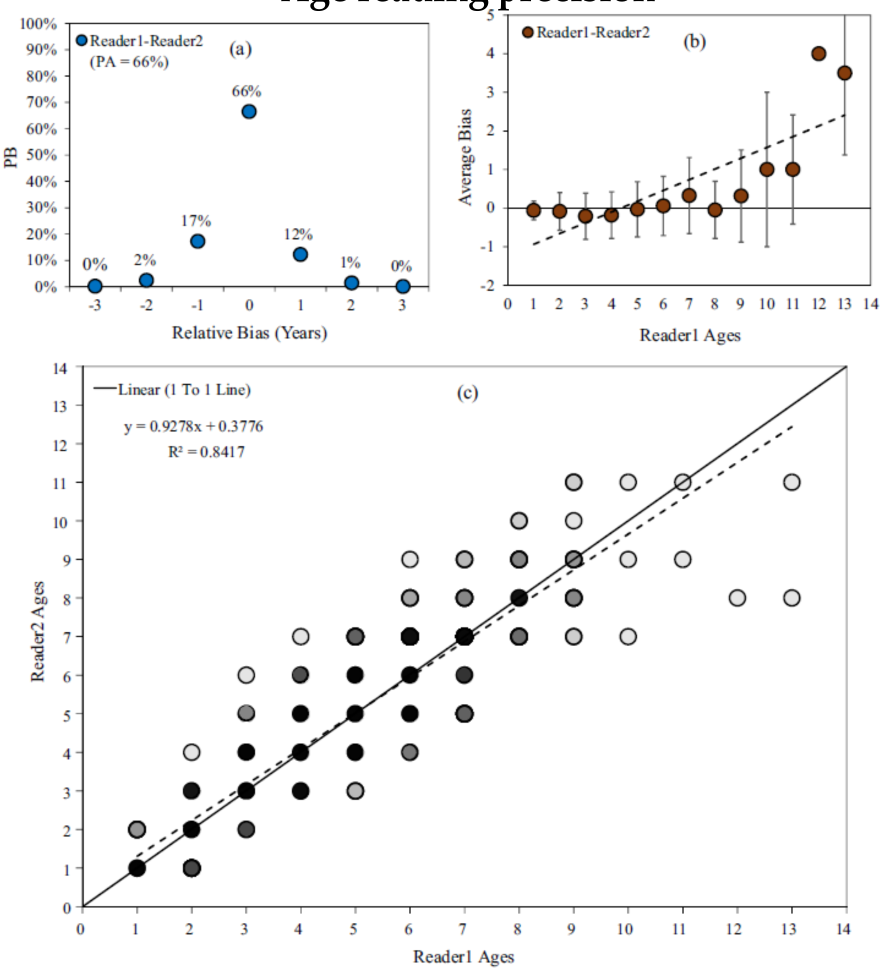


COASTAL AND MARINE ECOLOGY

HEALY ET AL.



Age reading precision



FT-NIRS ageing precision